THE FRONTIER EXASCALE SIMULATION
A LEAP FORWARD IN COSMOLOGICAL HYDRODYNAMICS

NICHOLAS FRONTIERE
CPS Division

mm Universe 2025
June 23rd 2025

Cosmological Simulations At Argonne

- **Role of Cosmology Simulations**
  - Simulations produce **full sky synthetic catalogs**, characterize systematics, and provide constraints of LCDM by solving the inverse problem
  - **Major impact** on multi-billion dollar DOE investment in ground and space-based sky surveys for planning, calibration, and validation
  - Requires **large-scale**, controlled accuracy, multi-physics simulations at high dynamic range
- **HACC (Hardware-Hybrid Accelerated Cosmology Code)**
  - **Hybrid** approach -- MPI + X (OpenMP, CUDA, HIP, SYCL, …)
  - Spectral Particle Mesh solver with a 3D FFT scalable to millions of ranks capable of processing grids **25000^3+ at scale**
  - Short range force solvers individually optimized to perform on both multi-core and **GPU accelerated** hardware
  - Capable of **extreme scale simulations** on all modern supercomputing hardware

LSST
Legacy Survey of Space and Time
WIDE-FIELD INFRARED SURVEY TELESCOPE
WFIRST
ASTROPHYSICS • DARK ENERGY • EXOPLANETS
DARK ENERGY SPECTROSCOPIC INSTRUMENT
DESI
SPHEREx
SOUTH POLE TELESCOPE
DES

Simulation Characteristics

Domain

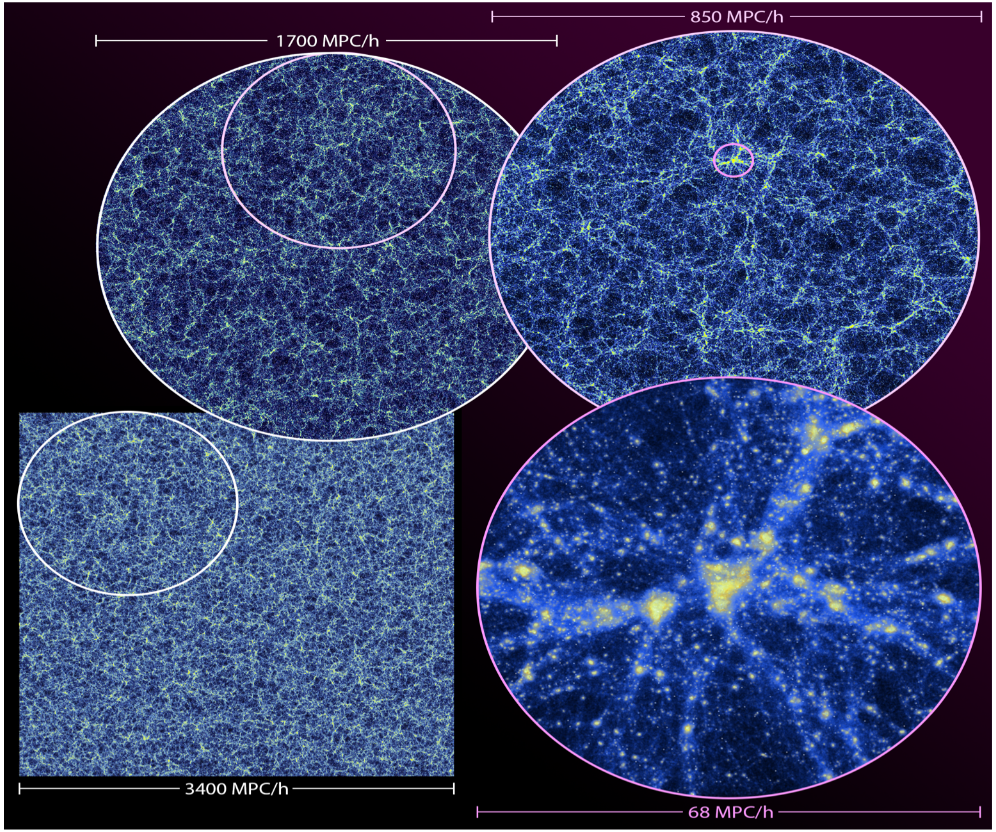
- 1-10 Gpc scale simulations
- Dynamic range of million to one
- Mass resolutions 10^7-10^9 $M_\odot$ h^{-1}

Gravity

- Vlasov-Poisson equation
- Dominates scales > ~1 Mpc
- N-body

Gas Physics

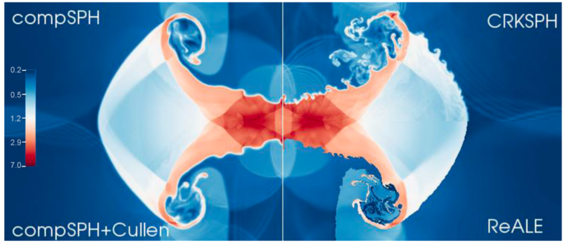
- High order SPH solver, CRKSPH
- UVB Heating and Atomic Cooling
- Supernovae and Star formation
- Chemical Enrichment
- AGN Feedback
- Emulated and calibrated to observations



Last Journey Simulation

CRKSPH Algorithm

- **Advantages of SPH**
  - Galilean invariant
  - Intrinsically adaptive
  - Agnostic to unique geometries
  - Conservative
  - Easily parallelizable/scalable

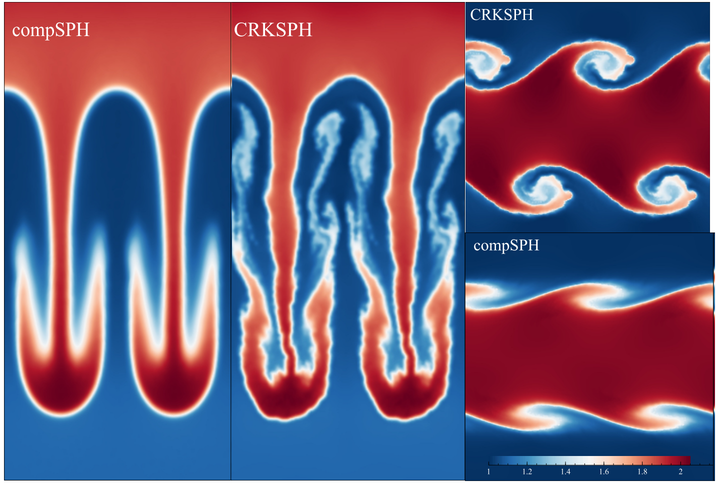


Triple-point shock

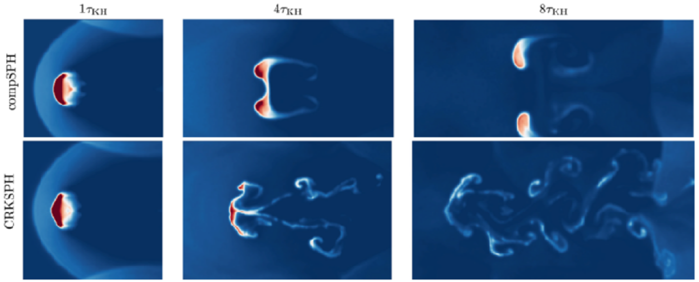
- **Disadvantages of (traditional) SPH**
  - Problems with shocks, gradients, and fluid mixing
  - Errors at zero'th order
  - Contact discontinuities violate formal assumptions
  - Problems with artificial viscosity
  - Many of these problems not removed at reasonable resolutions

- **Conservative Reproducing Kernel SPH**
  - Higher-order reproducing kernels exactly reproduce constant and linear order fields
  - Conservative reformulation of the dynamical equations maintains exact conservation of energy and momentum
  - New artificial viscosity that capitalizes on the increased accuracy of the velocity gradients, reduces excessive SPH artificial diffusion
  - No artificial conductivity required

Rayleigh-Taylor **Kelvin-Helmholtz**



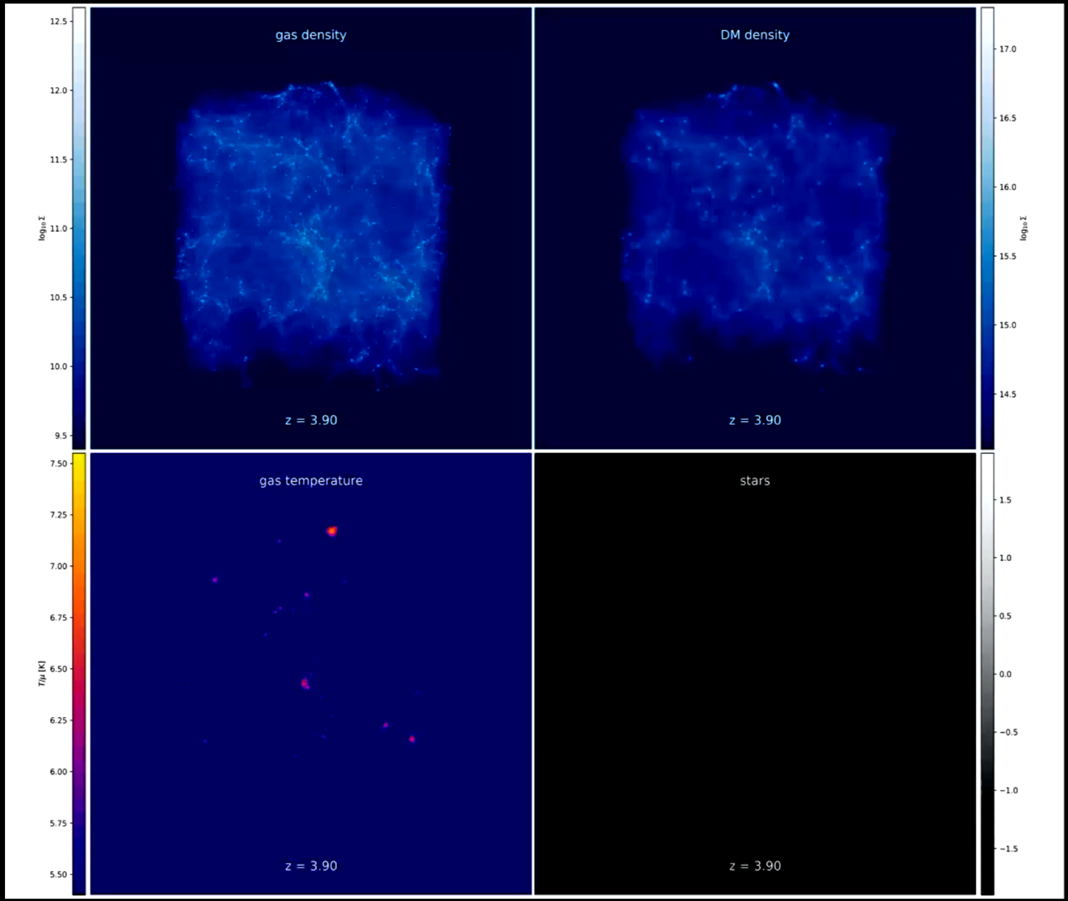
Frontiere et al., J. Comp. Phys. 332, 160 (2017)



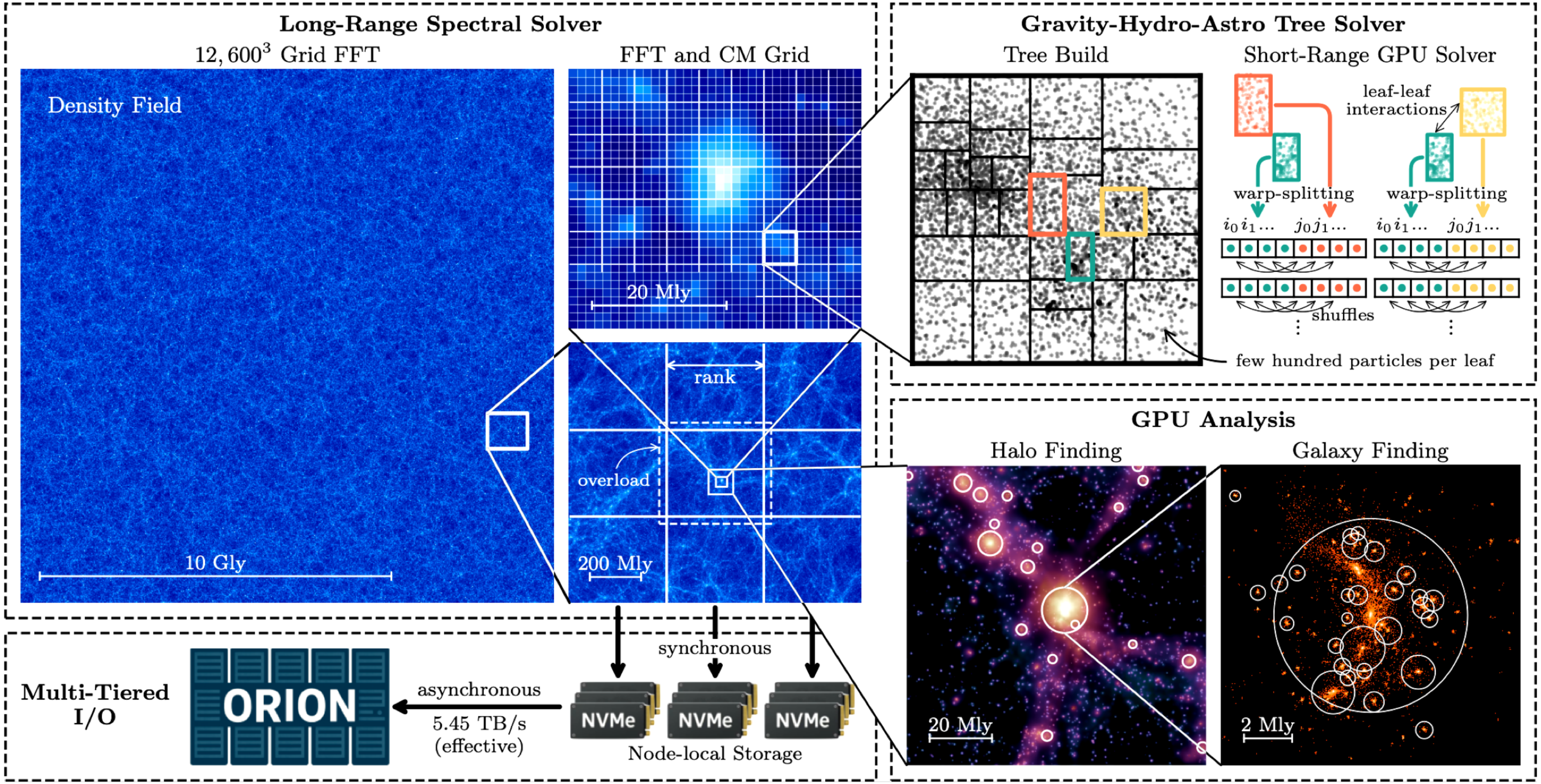
'Blob' test with CRK-SPH showing blob disruption due to Kelvin-Helmholtz and Rayleigh-Taylor instabilities

Multi-physics Coupling

Multi-physics Coupling



CRK-HACC Runtime Overview



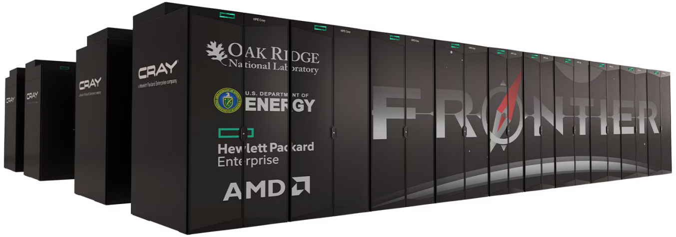
100 PB of data written < 3% TTS

All compute intensive solvers and analysis are offloaded to the **GPU**

At scale on Frontier achieves **513 PFLOPs** of performance evaluating **46.6 billion** particles/s

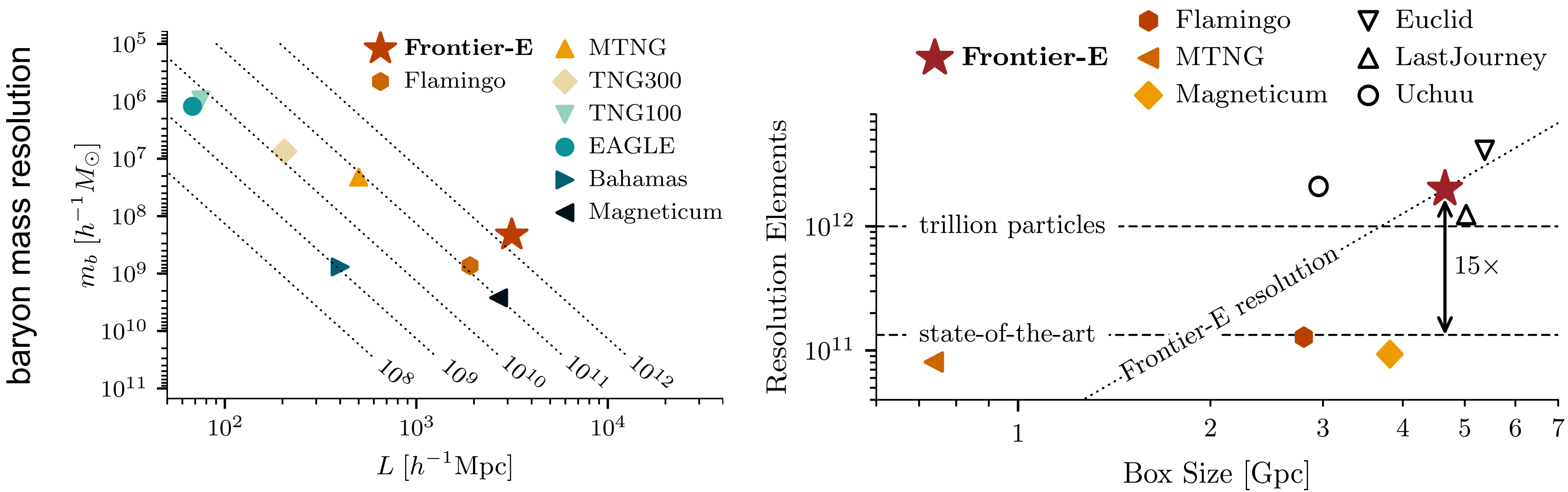
Age of Exascale

- **Next generation of supercomputers**
  - Frontier ORNL 1.353 Eflop/s
  - Aurora ANL 1.012 Eflop/s
  - El Capitan LLNL 1.742 Eflop/s
  - GPU accelerated
  - Mixed precision



- **Exascale class capabilities**
  - 10-50x improvement
  - Hydro simulations are typically 10-30x slower than GO
  - Previous state-of-the-art GO simulations can now include gas physics. Trillion(s) particles
  - Next-generation GO simulations in the 10+ trillion particle class.

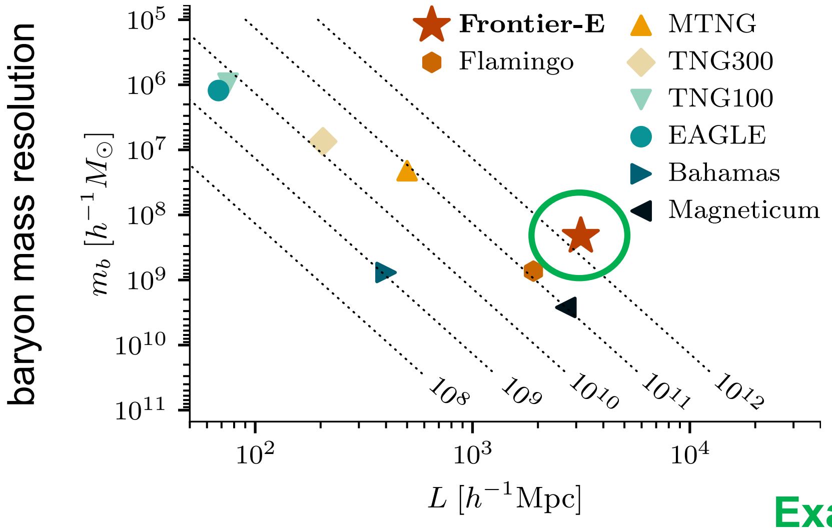
Exascale Class Simulations

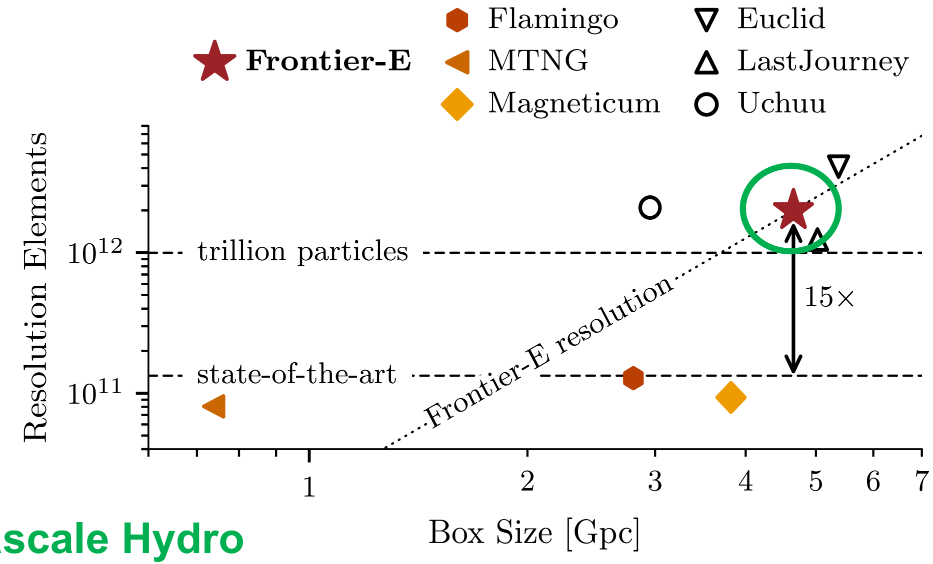


Exascale unlocks the ability to run previous **state of the art** gravity only simulations now including **hydrodynamics** in a **week** of machine time (~ 1 yr CPU)

Gravity only simulations break into the **10+ trillion** particle class

Exascale Class Simulations



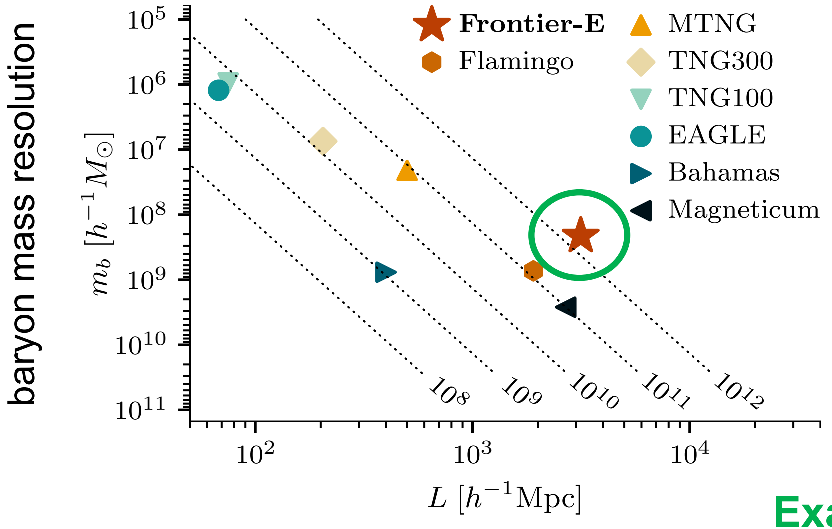


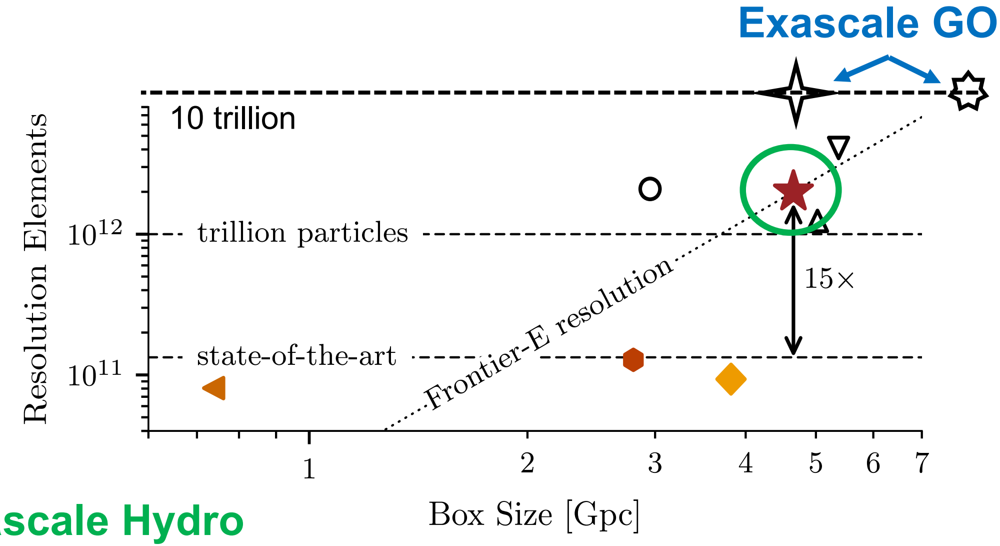
Exascale Hydro

Exascale unlocks the ability to run previous **state of the art** gravity only simulations now including **hydrodynamics** in a **week** of machine time (~ 1 yr CPU)

Gravity only simulations break into the **10+ trillion** particle class

Exascale Class Simulations





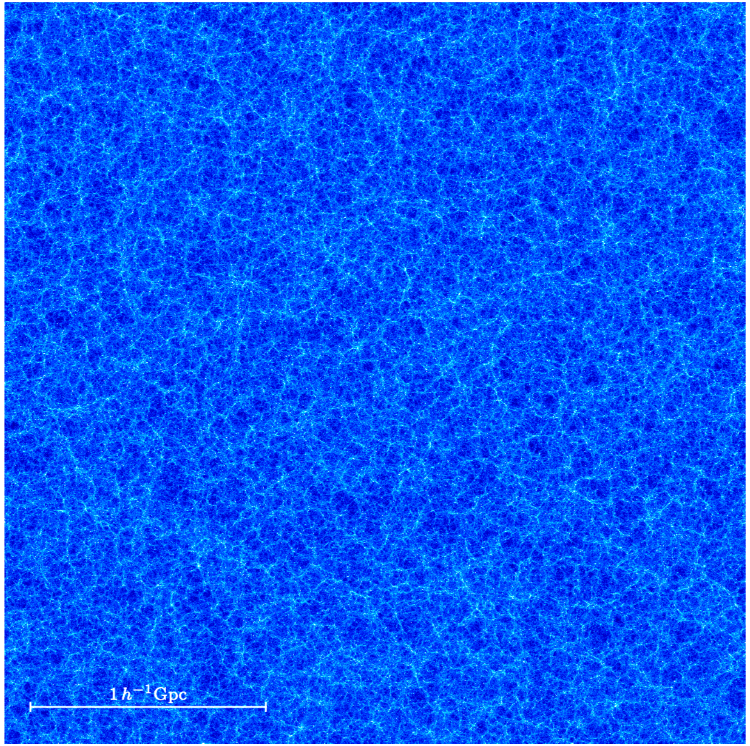
Exascale Hydro

Exascale unlocks the ability to run previous **state of the art** gravity only simulations now including **hydrodynamics** in a **week** of machine time (~ 1 yr CPU)

Gravity only simulations break into the **10+ trillion** particle class

Frontier Exascale Simulation

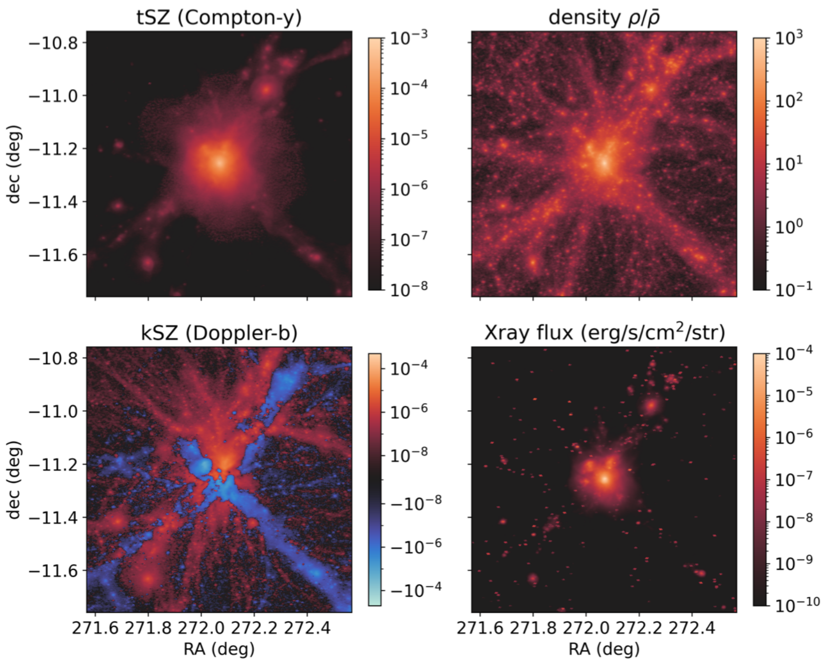
- **Survey Scale**
  - 3.150 Gpc h^{-1} domain
  - 2 x 12600^3 particles – 4 trillion
  - 2 x 10^8 $M_\odot$ h^{-1} baryon mass resolution
- **Cluster Science**
  - 2000+ (13) clusters > 10^{15} $M_\odot$ h^{-1}
  - 570,000+ (75k) > 10^{14} $M_\odot$ h^{-1}
- **Full sky particle LC (z < 3)**
  - Consistent maps and catalogs
  - Downscale LC out to z < 5
- **Extensive Data Outputs (12 PBs)**
  - Full particles for objects > 10^{13} $M_\odot$ h^{-1}
  - Merger trees, catalogs and snapshots
  - ~ 1 billion galaxies at z = 0



Evolved to z = 0, Frontier-E marks a new generation of survey-scale simulations with over an order-of-magnitude leap in scale

Frontier Exascale Simulation

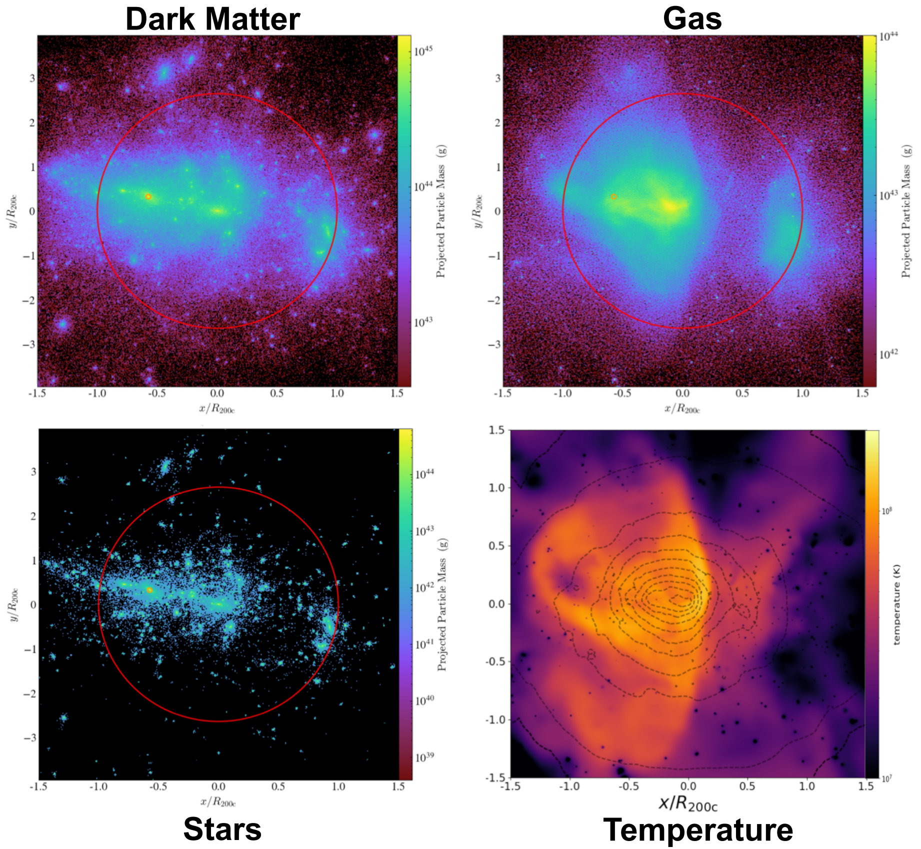
- **Survey Scale**
  - 3.150 Gpc h^{-1} domain
  - 2 x 12600^3 particles – 4 trillion
  - 2 x 10^8 $M_\odot$ h^{-1} baryon mass resolution
- **Cluster Science**
  - 2000+ (13) clusters > 10^{15} $M_\odot$ h^{-1}
  - 570,000+ (75k) > 10^{14} $M_\odot$ h^{-1}
- **Full sky particle LC (z < 3)**
  - Consistent maps and catalogs
  - Downscale LC out to z < 5
- **Extensive Data Outputs (12 PBs)**
  - Full particles for objects > 10^{13} $M_\odot$ h^{-1}
  - Merger trees, catalogs and snapshots
  - ~ 1 billion galaxies at z = 0



Mock observations of a cluster on the light cone. Thermal and kinematic Sunyaev-Zeldovich maps, the total matter density contrast, and the flux emission in a bolometric X-ray band

Frontier Exascale Simulation

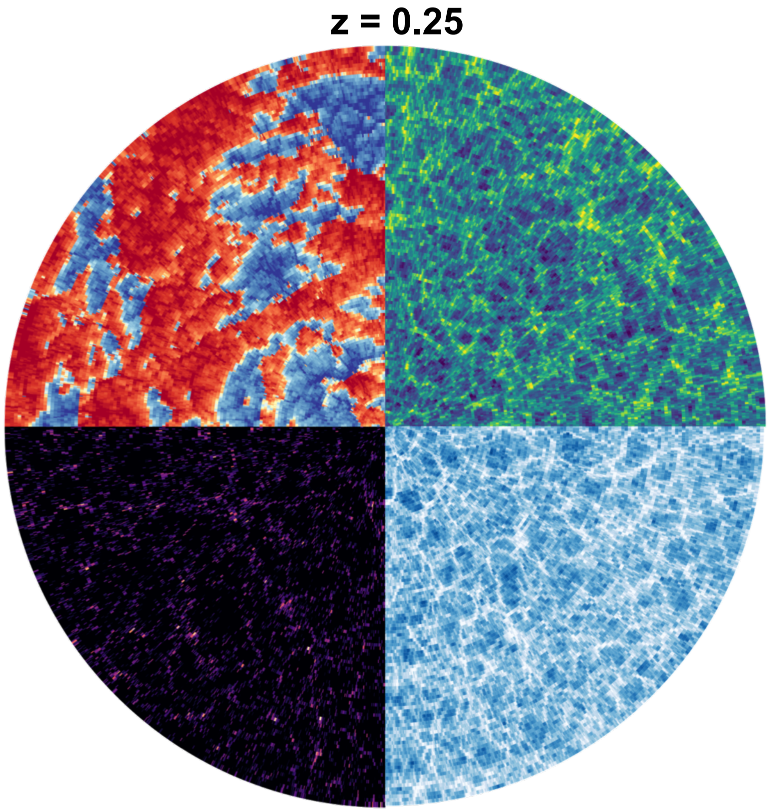
- **Survey Scale**
  - 3.150 Gpc h^{-1} domain
  - 2 x 12600^3 particles – 4 trillion
  - 2 x 10^8 $M_\odot$ h^{-1} baryon mass resolution
- **Cluster Science**
  - 2000+ (13) clusters > 10^{15} $M_\odot$ h^{-1}
  - 570,000+ (75k) > 10^{14} $M_\odot$ h^{-1}
- **Full sky particle LC (z < 3)**
  - Consistent maps and catalogs
  - Downscale LC out to z < 5
- **Extensive Data Outputs (12 PBs)**
  - Full particles for objects > 10^{13} $M_\odot$ h^{-1}
  - Merger trees, catalogs and snapshots
  - ~ 1 billion galaxies at z = 0



The large volume enables the identification and study of rare, dynamically complex systems

Frontier Exascale Simulation

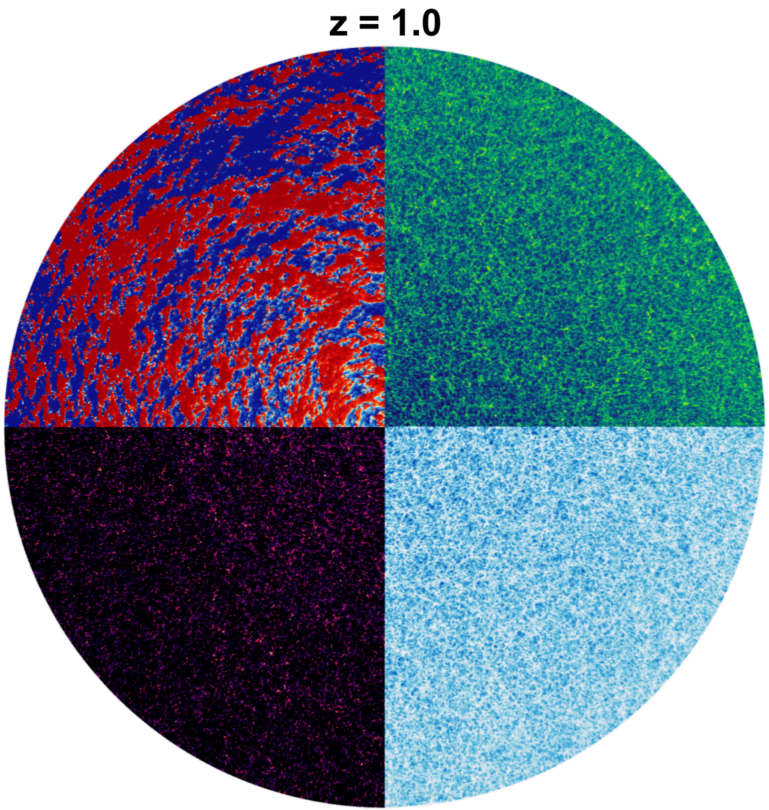
- **Survey Scale**
  - 3.150 Gpc h^{-1} domain
  - 2 x 12600^3 particles – 4 trillion
  - 2 x 10^8 $M_\odot h^{-1}$ baryon mass resolution
- **Cluster Science**
  - 2000+ (13) clusters > 10^{15} $M_\odot h^{-1}$
  - 570,000+ (75k) > 10^{14} $M_\odot h^{-1}$
- **Full sky particle LC (z < 3)**
  - Consistent maps and catalogs
  - Downscale LC out to z < 5
- **Extensive Data Outputs (12 PBs)**
  - Full particles for objects > 10^{13} $M_\odot h^{-1}$
  - Merger trees, catalogs and snapshots
  - ~ 1 billion galaxies at z = 0



Slice through full-sky maps showing, clockwise from top left: kSZ, tSZ, X-ray, and density.

Frontier Exascale Simulation

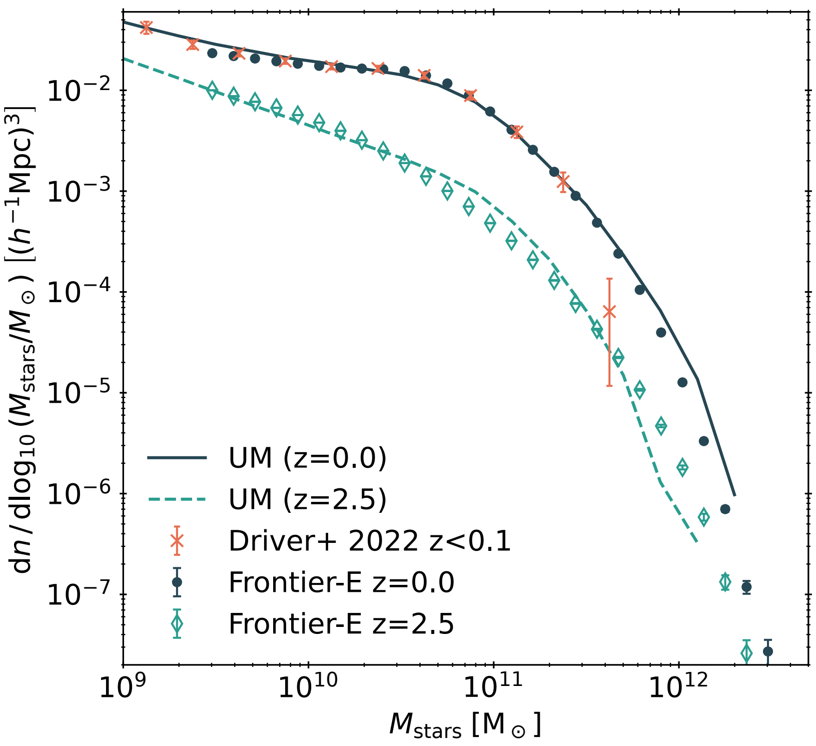
- **Survey Scale**
  - 3.150 Gpc h^{-1} domain
  - 2 x 12600^3 particles – 4 trillion
  - 2 x 10^8 $M_\odot$ h^{-1} baryon mass resolution
- **Cluster Science**
  - 2000+ (13) clusters > 10^{15} $M_\odot$ h^{-1}
  - 570,000+ (75k) > 10^{14} $M_\odot$ h^{-1}
- **Full sky particle LC (z < 3)**
  - Consistent maps and catalogs
  - Downscale LC out to z < 5
- **Extensive Data Outputs (12 PBs)**
  - Full particles for objects > 10^{13} $M_\odot$ h^{-1}
  - Merger trees, catalogs and snapshots
  - ~ 1 billion galaxies at z = 0



Slice through full-sky maps showing, clockwise from top left: kSZ, tSZ, X-ray, and density.

Frontier Exascale Simulation

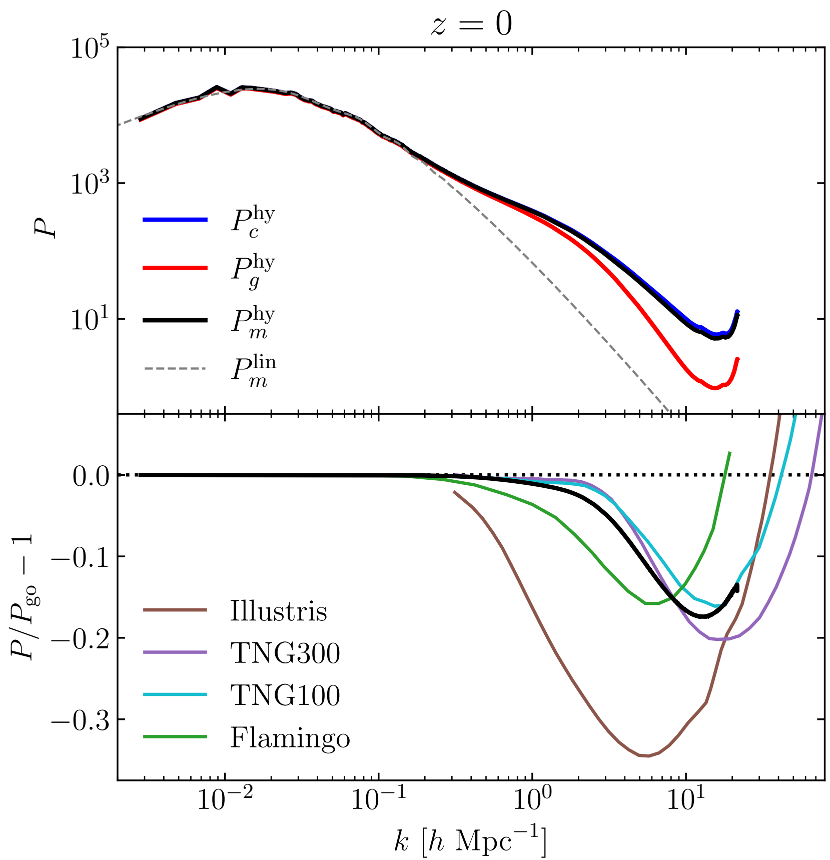
- **Survey Scale**
  - 3.150 Gpc h^{-1} domain
  - 2 x 12600^3 particles – 4 trillion
  - 2 x 10^8 $M_\odot$ h^{-1} baryon mass resolution
- **Cluster Science**
  - 2000+ (13) clusters > 10^{15} $M_\odot$ h^{-1}
  - 570,000+ (75k) > 10^{14} $M_\odot$ h^{-1}
- **Full sky particle LC (z < 3)**
  - Consistent maps and catalogs
  - Downscale LC out to z < 5
- **Extensive Data Outputs (12 PBs)**
  - Full particles for objects > 10^{13} $M_\odot$ h^{-1}
  - Merger trees, catalogs and snapshots
  - ~ 1 billion galaxies at z = 0



Calibrated Galaxy Stellar Mass Function

Frontier Exascale Simulation

- **Survey Scale**
  - 3.150 Gpc h^{-1} domain
  - 2 x 12600^3 particles – 4 trillion
  - 2 x 10^8 $M_\odot$ h^{-1} baryon mass resolution
- **Cluster Science**
  - 2000+ (13) clusters > 10^{15} $M_\odot$ h^{-1}
  - 570,000+ (75k) > 10^{14} $M_\odot$ h^{-1}
- **Full sky particle LC (z < 3)**
  - Consistent maps and catalogs
  - Downscale LC out to z < 5
- **Extensive Data Outputs (12 PBs)**
  - Full particles for objects > 10^{13} $M_\odot$ h^{-1}
  - Merger trees, catalogs and snapshots
  - ~ 1 billion galaxies at z = 0



Power Spectrum Comparisons

Ember Cluster

1/100,000th subvolume largest cluster zoom-in

Ember Cluster

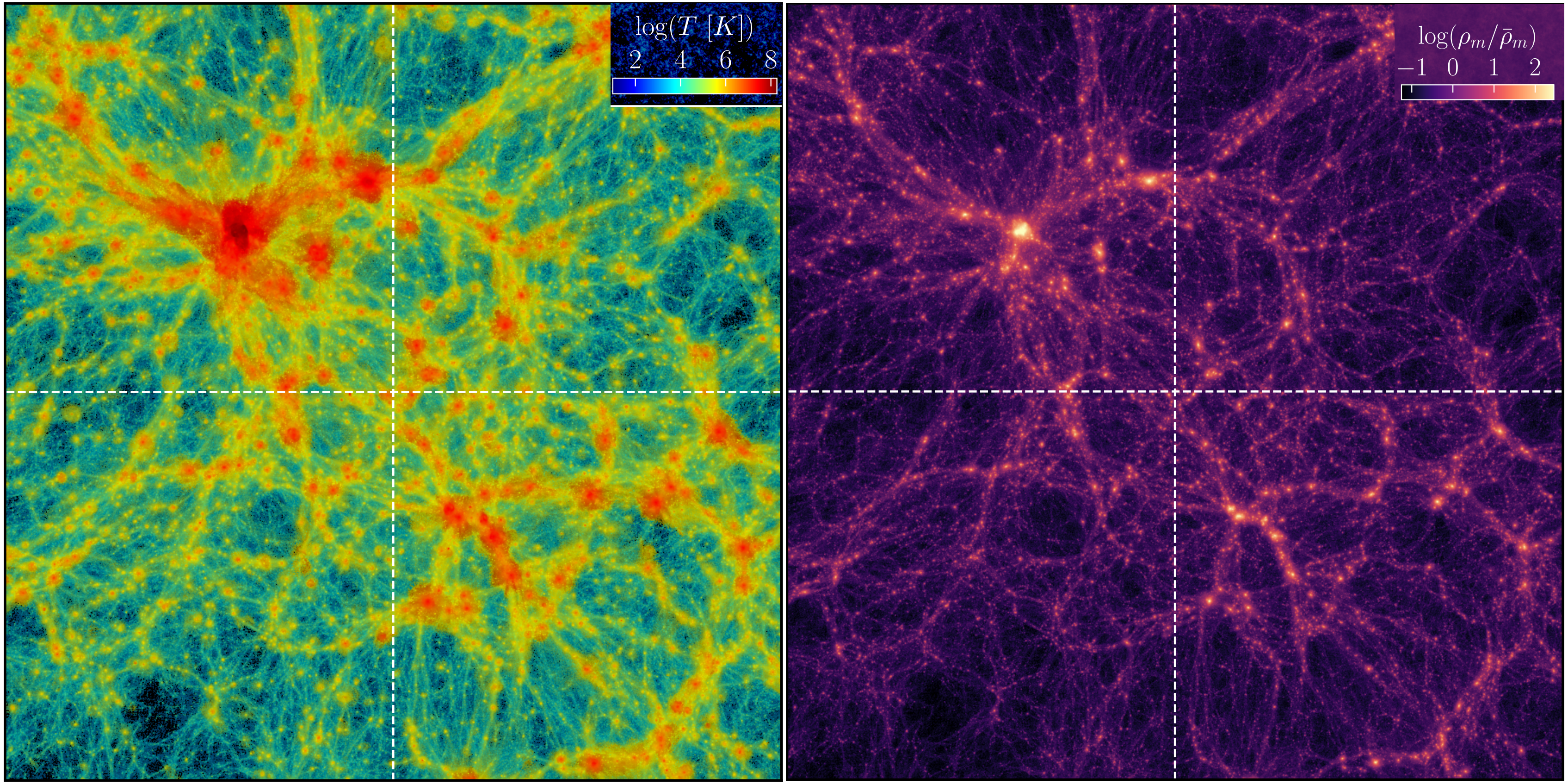
Frontier-E Simulation

N. Frontiere[1], M. Buehlmann[1], J.D. Emberson[1],
S. Habib[1,2], P. Larsen[1], K. Heitmann[2], C.-A. Faucher-Giguère[3]

[1]CPS Division, Argonne National Laboratory, Lemont, IL 60439, USA
[2]HEP Division, Argonne National Laboratory, Lemont, IL 60439, USA
[3]CIERA and Department of Physics and Astronomy, Northwestern University, Evanston, IL 60201, USA

1/100,000th subvolume largest cluster zoom-in

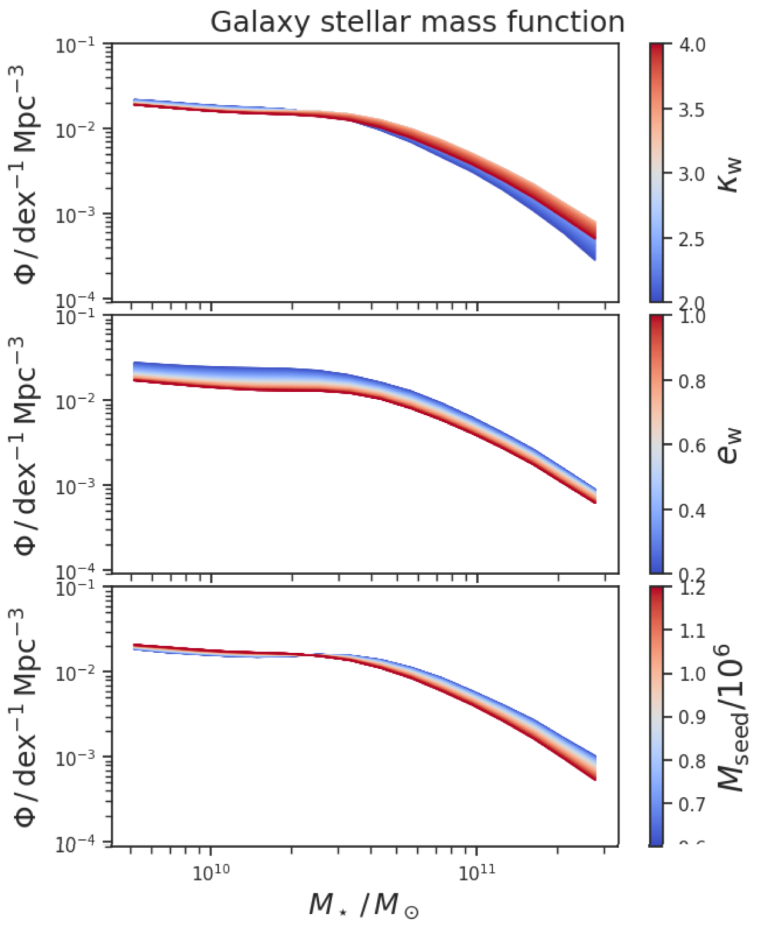
Frontier Exascale Simulation



Four ranks out of 72,000

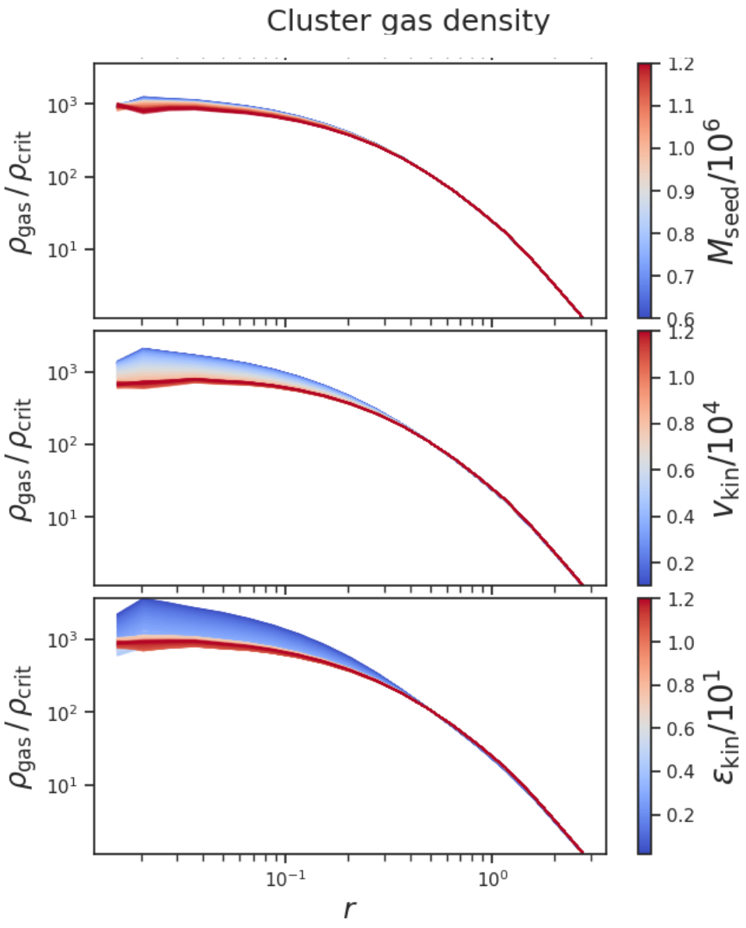
Frontier Exascale Calibration

- **Simulation Suites**
  - 128 Mpc h^{-1} initial calibration set
  - 256 Mpc h^{-1} cluster calibration
- **Targets**
  - Multi-redshift Galaxy Stellar Mass Function
  - Cluster Gas Density Profiles
- **Emulation**
  - Varying across 5 active parameters
  - Two wind, three for AGN
- **Exascale Throughput**
  - Exploration at larger volumes
  - Parameter expansion – hydro + cosmology



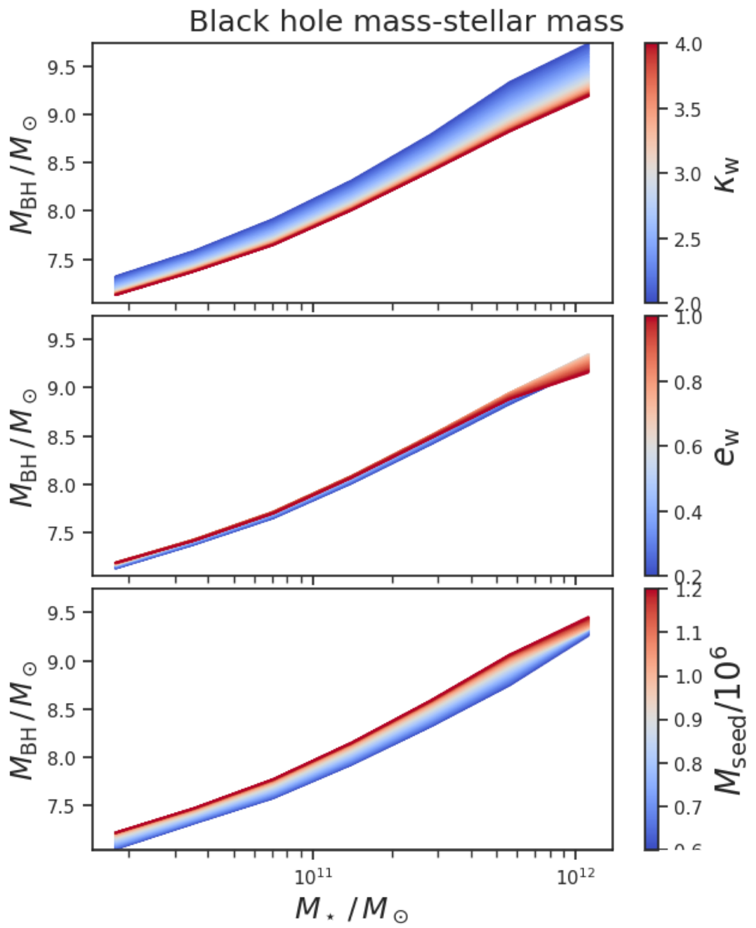
Frontier Exascale Calibration

- **Simulation Suites**
  - 128 Mpc h^{-1} initial calibration set
  - 256 Mpc h^{-1} cluster calibration
- **Targets**
  - Multi-redshift Galaxy Stellar Mass Function
  - Cluster Gas Density Profiles
- **Emulation**
  - Varying across 5 active parameters
  - Two wind, three for AGN
- **Exascale Throughput**
  - Exploration at larger volumes
  - Parameter expansion – hydro + cosmology



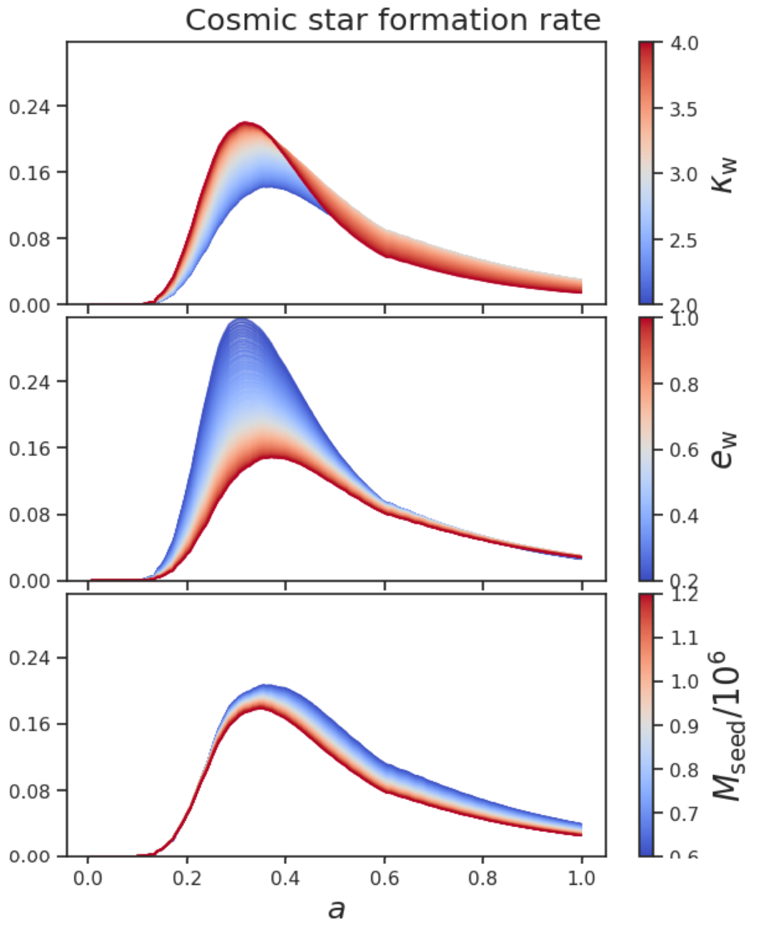
Frontier Exascale Calibration

- **Simulation Suites**
  - 128 Mpc h^{-1} initial calibration set
  - 256 Mpc h^{-1} cluster calibration
- **Targets**
  - Multi-redshift Galaxy Stellar Mass Function
  - Cluster Gas Density Profiles
- **Emulation**
  - Varying across 5 active parameters
  - Two wind, three for AGN
- **Exascale Throughput**
  - Exploration at larger volumes
  - Parameter expansion – hydro + cosmology



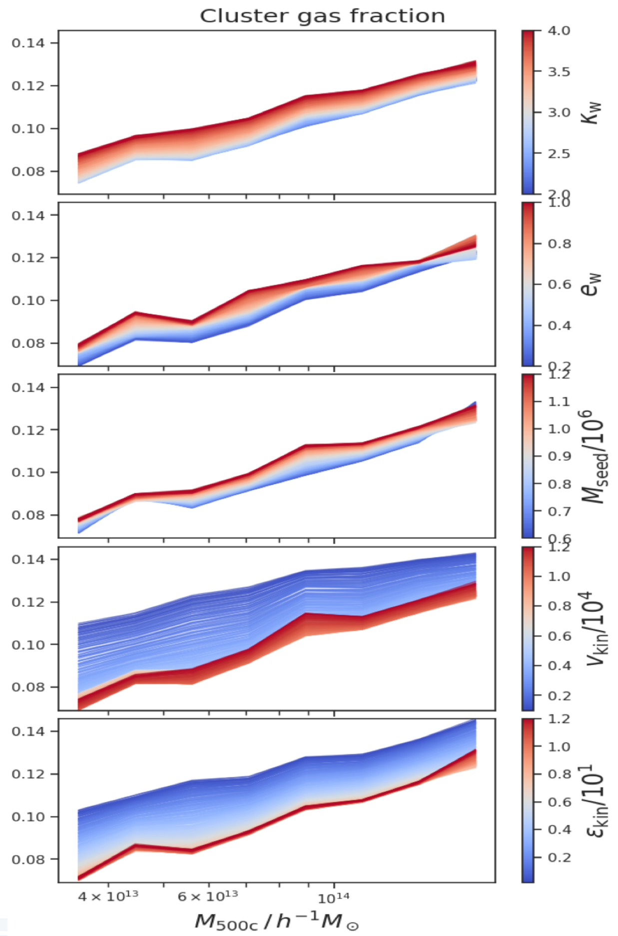
Frontier Exascale Calibration

- **Simulation Suites**
  - 128 Mpc h^{-1} initial calibration set
  - 256 Mpc h^{-1} cluster calibration
- **Targets**
  - Multi-redshift Galaxy Stellar Mass Function
  - Cluster Gas Density Profiles
- **Emulation**
  - Varying across 5 active parameters
  - Two wind, three for AGN
- **Exascale Throughput**
  - Exploration at larger volumes
  - Parameter expansion – hydro + cosmology



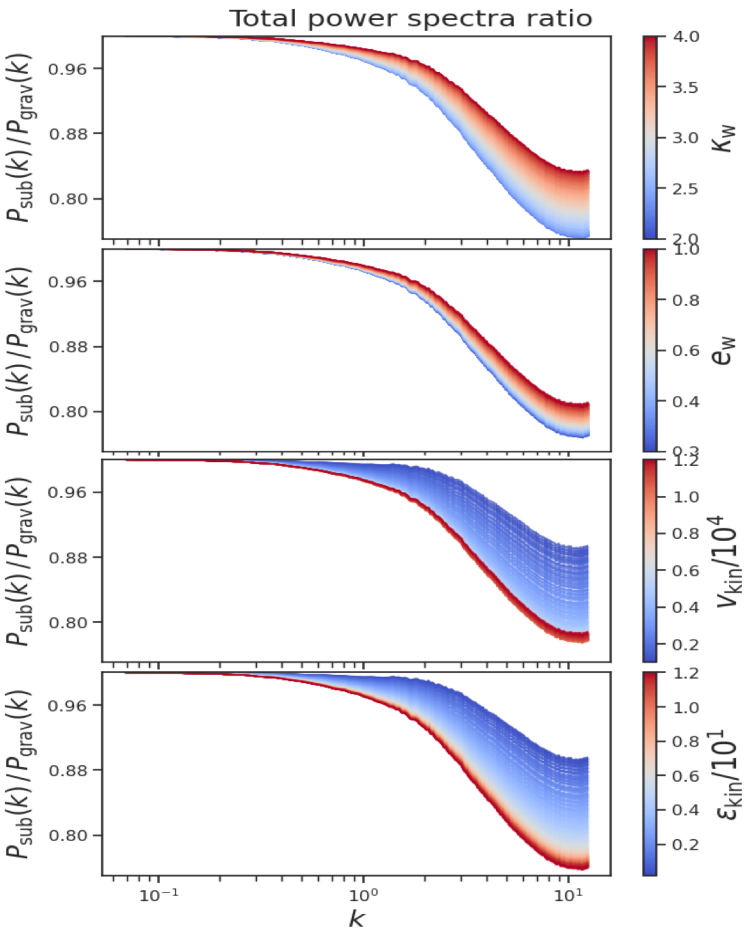
Frontier Exascale Calibration

- **Simulation Suites**
  - 128 Mpc h^{-1} initial calibration set
  - 256 Mpc h^{-1} cluster calibration
- **Targets**
  - Multi-redshift Galaxy Stellar Mass Function
  - Cluster Gas Density Profiles
- **Emulation**
  - Varying across 5 active parameters
  - Two wind, three for AGN
- **Exascale Throughput**
  - Exploration at larger volumes
  - Parameter expansion – hydro + cosmology



Frontier Exascale Calibration

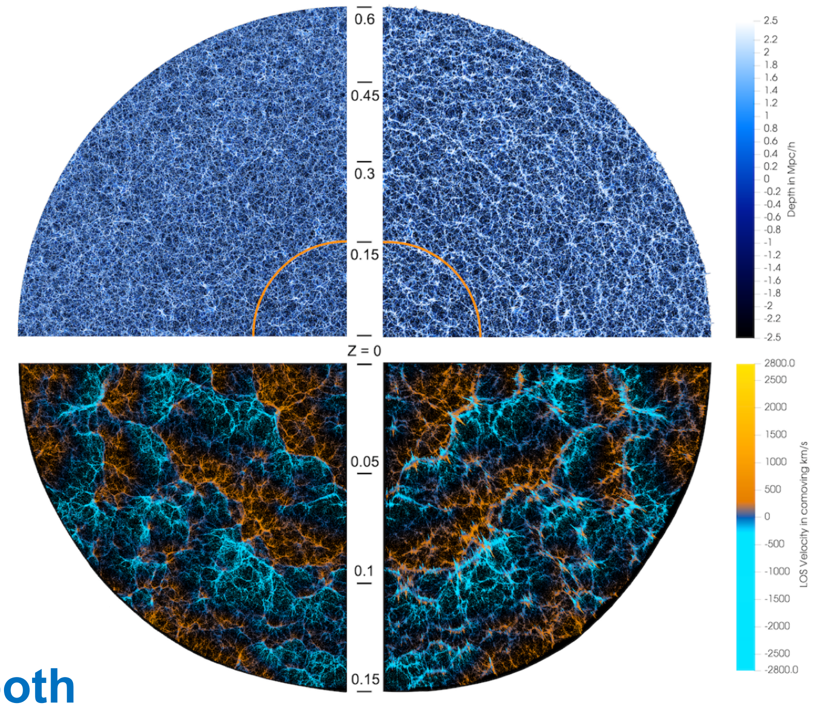
- **Simulation Suites**
  - 128 Mpc h^{-1} initial calibration set
  - 256 Mpc h^{-1} cluster calibration
- **Targets**
  - Multi-redshift Galaxy Stellar Mass Function
  - Cluster Gas Density Profiles
- **Emulation**
  - Varying across 5 active parameters
  - Two wind, three for AGN
- **Exascale Throughput**
  - Exploration at larger volumes
  - Parameter expansion – hydro + cosmology



Exascale Class Gravity Simulations

Upcoming Aurora Exascale Simulation Pair

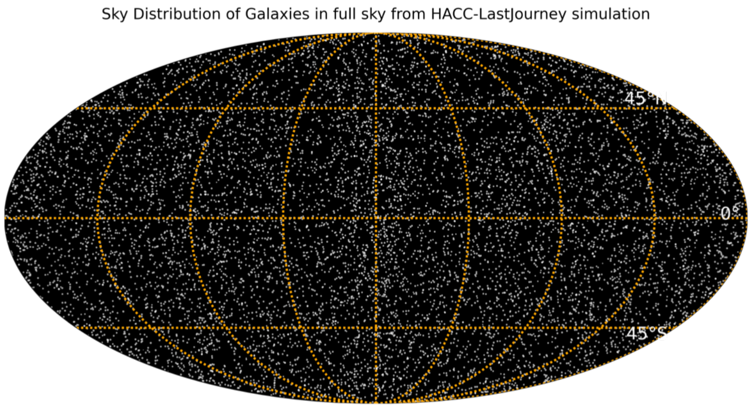
- **Primordial non-Gaussianity GO suite**
  - 5.5 Gpc/h domain
  - $22{,}000^3$ particles – **10 trillion**
  - $10^9\ M_\odot h^{-1}$ mass resolution
- **Investigate PNG in $1 > k > 10^{-3}$ Mpc^{-1}**
  - Blinded analyses
  - SphereX full sky and DESI mocks
  - Assembly histories used for simulated sky observables
- **Synthetic sky catalogs generated using both galaxy-halo connection and ML methods**

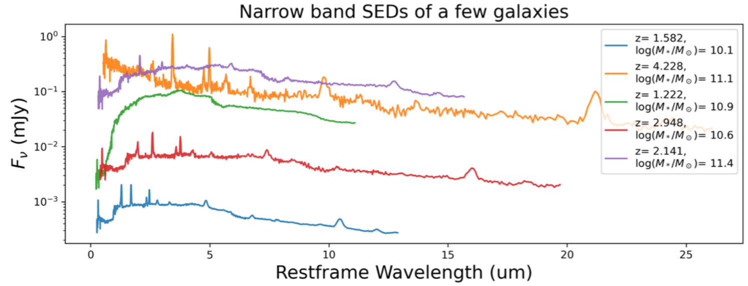


Particle light-cone visualizations from the 'Farpoint' run (Frontiere et al., ApJS 2022), a high mass resolution, large-volume cosmological simulation run with HACC

Machine Learning Analyses

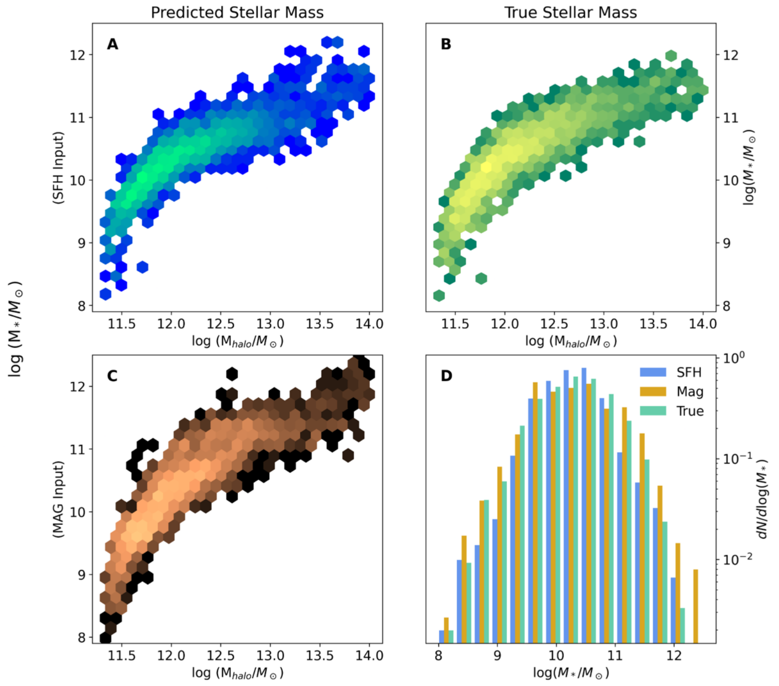
- **SED Painting N-body Simulations**
  - Mass accretion histories, are connected with star formation histories and dust attenuation curves
  - Produces billions of galaxies covering full sky, across survey sensitive redshift ranges
- **Painting galaxies with colors**
  - Input stellar age, metallicity and SSP spectral emissions
  - Output catalogs include luminosity profiles, galaxy SEDs along with photometry for Rubin, SPHEREx, COSMOS and other surveys
- **Multi-modal, multi-task model training**
  - Scalar catalogs, time-series, SEDs, merger trees, galaxy images, density fields: multiple modalities representing multiple representations of the same object
  - Foundation model connects simulation values with and observables





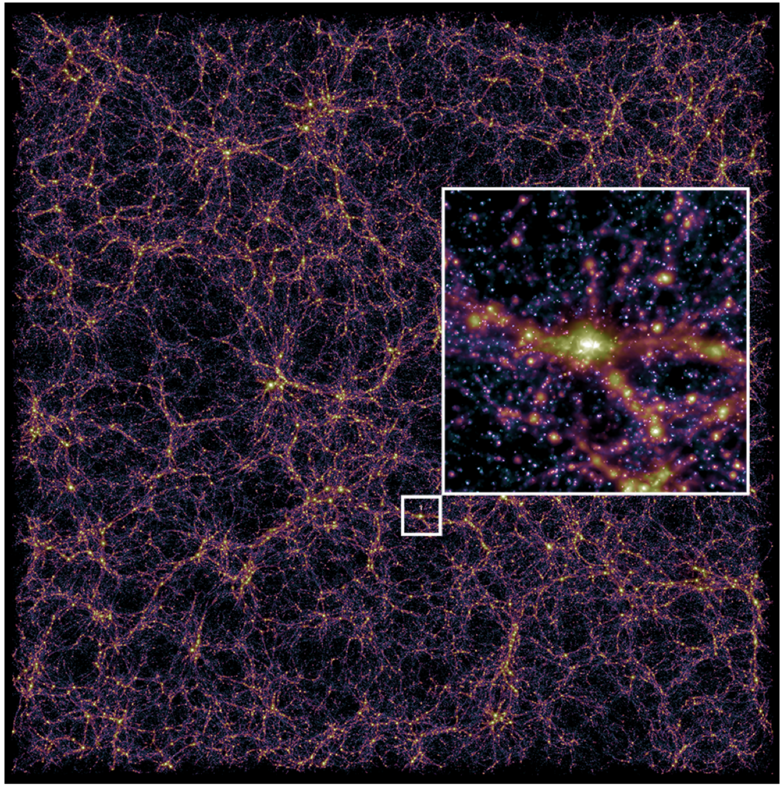
Machine Learning Analyses

- **SED Painting N-body Simulations**
  - Mass accretion histories, are connected with star formation histories and dust attenuation curves
  - Produces billions of galaxies covering full sky, across survey sensitive redshift ranges
- **Painting galaxies with colors**
  - Input stellar age, metallicity and SSP spectral emissions
  - Output catalogs include luminosity profiles, galaxy SEDs along with photometry for Rubin, SPHEREx, COSMOS and other surveys
- **Multi-modal, multi-task model training**
  - Scalar catalogs, time-series, SEDs, merger trees, galaxy images, density fields: multiple modalities representing multiple representations of the same object
  - Foundation model connects simulation values with and observables



The Exascale Future is Bright

- Exascale machines provide a **powerful tool** to drive numerical cosmology science
- Enabling **>10× leap** in simulation scale: from trillion-particle hydro to 10+ trillion GO in ~week(s) of machine time.
- Generated datasets are massive, ideal for **ML training**
- Capable of running targeted **emulation suites** in new optimization domains
- Will significantly advance **predictive capabilities** and **simulation fidelity** necessary to validate and interpret results from ongoing and upcoming surveys



slice through downscaled Frontier-E simulation raytracing visualization

QUESTIONS?

EXTRA SLIDES (PLOTS)

CRK-HACC: GPU-Accelerated Analysis

DBSCAN

- Density-Based Spatial Clustering of Applications with Noise
- Collect particles that are within a “linking length” of each other
- Define the minimum number of neighbor points (MinPts)

FOF Finder

- “Friends-of-Friends”
- Commonly used to find dark matter halos
- FOF is DBSCAN with MinPts = 1

SOD

- Spherical Overdensity Halo finding
- Integrated spherical shells to a given mass threshold
- Used for both SO halo catalogs as well as galaxy catalogs

ArborX

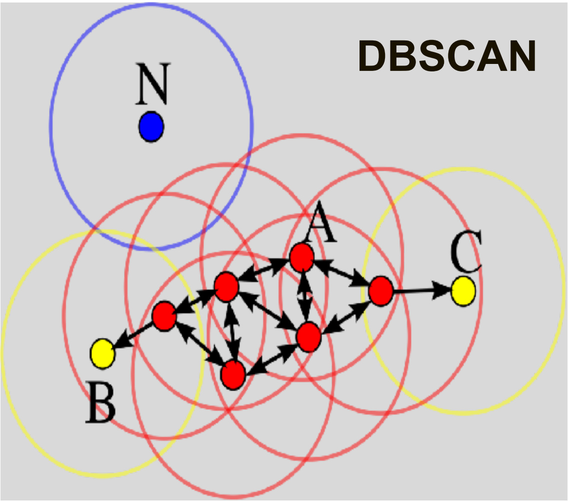
- Implemented DBSCAN on top of BVH (Bounding Volume Hierarchy) tree and optimized performance
- Publicly available at https://github.com/arborx/ArborX/wiki/DBSCAN

Performance

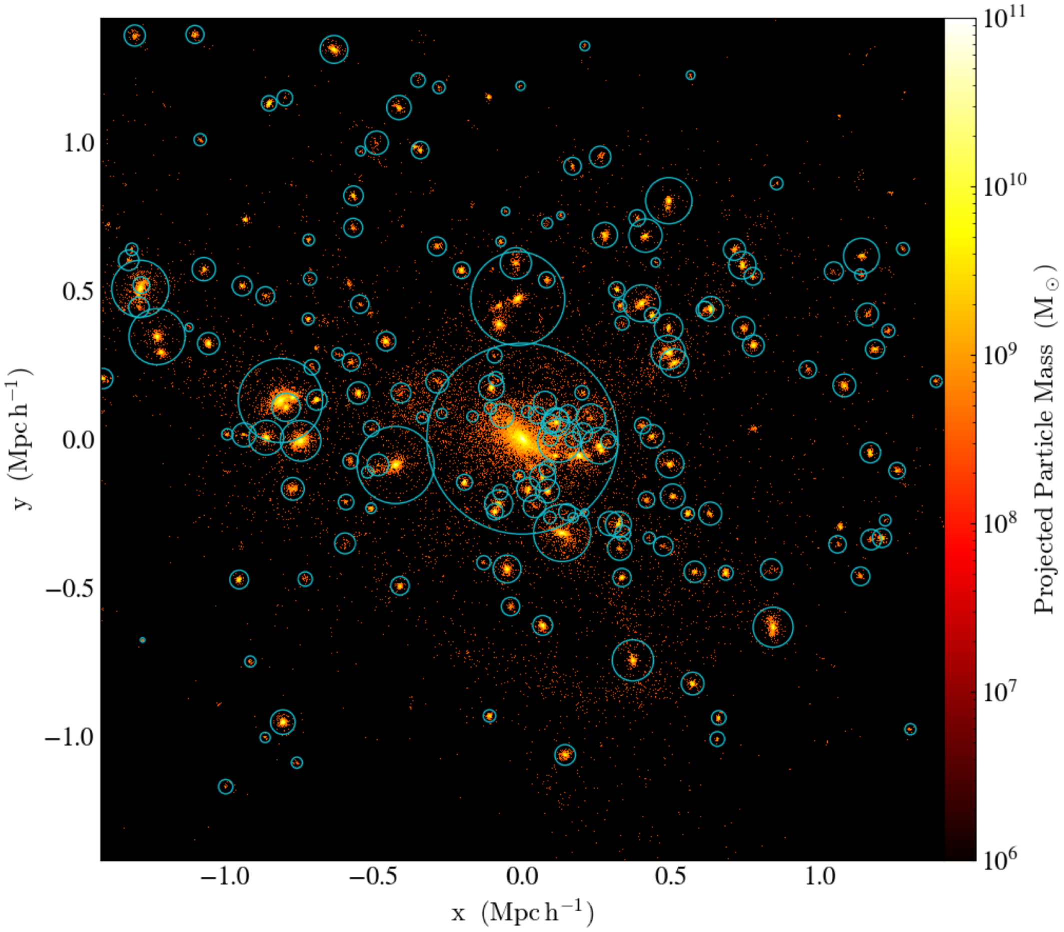
- The ArborX implementation is a factor of **10x faster** than the optimized OpenMP threaded CPU version of FOF finding

Hierarchical DBSCAN

- Future work replacement for “subhalo” finding

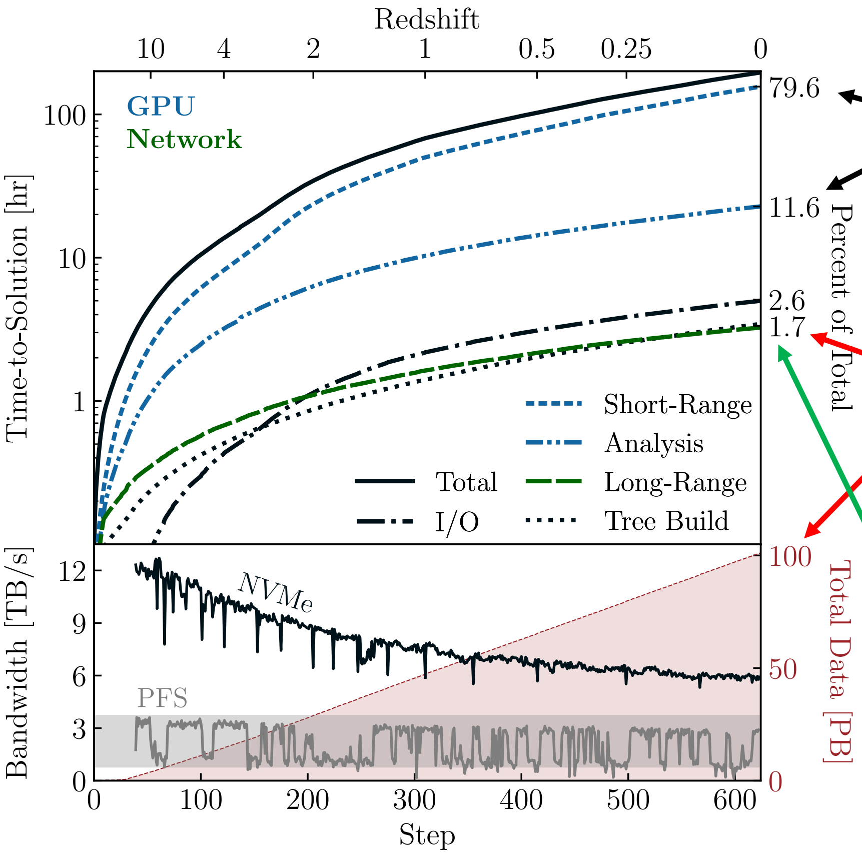


CRK-HACC: Galaxy Finding with GPU accelerated DBSCAN



Frontier-E Cumulative Time to Solution

- Short-range computational operations dominate TTS
- Asynchronous data bleeding using NVMe provides high-bandwidth I/O
- Complete in 1 week of machine time. On the CPUs of Frontier, the same simulation would take ~1 year to run

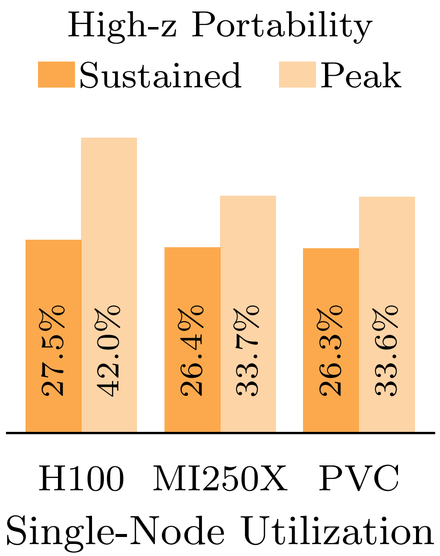


Over 90% TTS was performed on the GPU

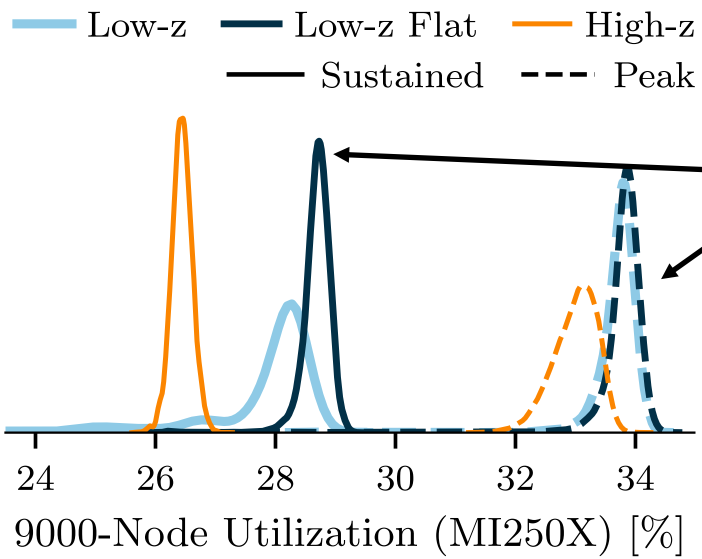
I/O was 2.6% despite writing over 100 PB of data

Long-range solver subdominant despite running FFTs on more than two trillion cells

Frontier-E GPU Efficiency



Single-node performance similar across GPU vendors

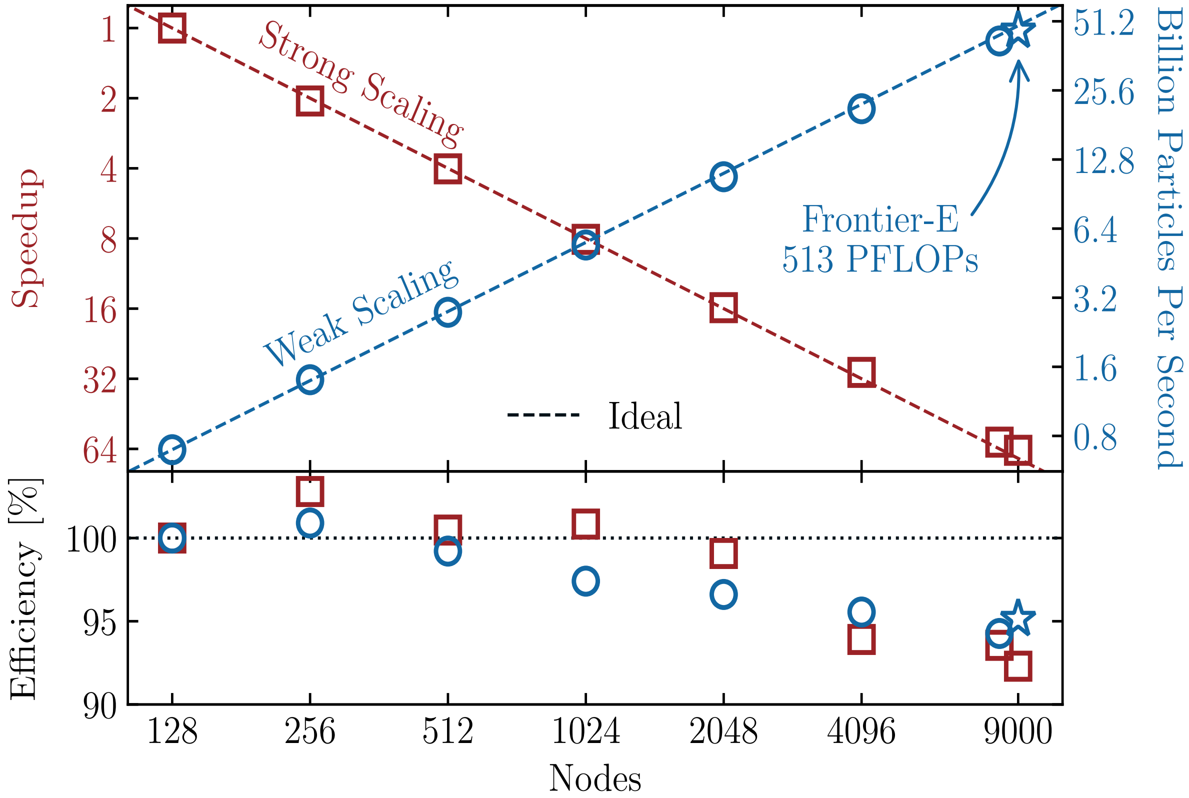


9000-node per-GPU performance distribution at high and low redshift

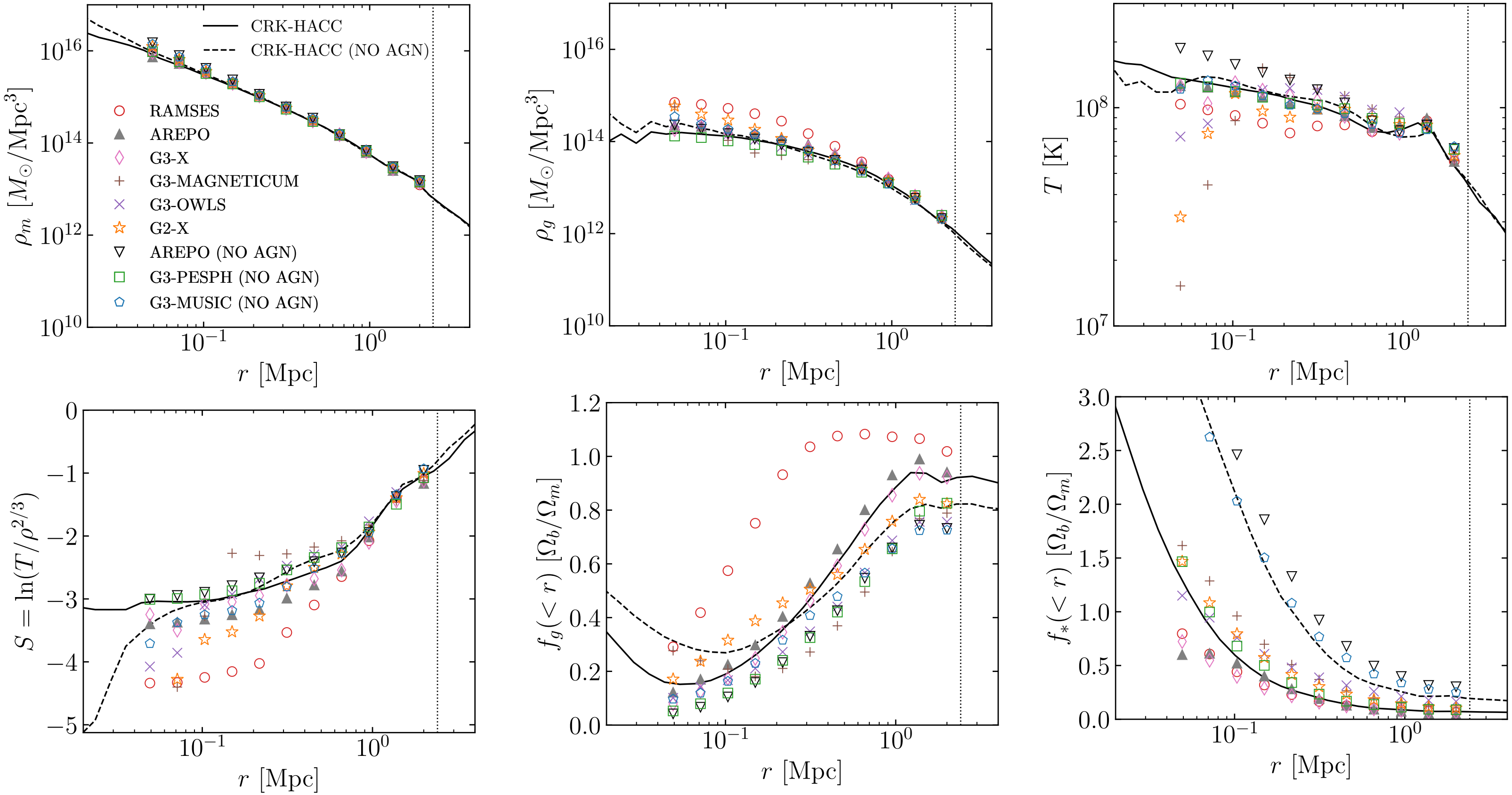
Flat vs. adaptive time-stepper maintains peak and sustained utilization

Frontier-E Scaling

- Near-ideal weak and strong scaling
- Achieved 513 PFLOPs on full machine
- Processing 46.6 billion particles per second

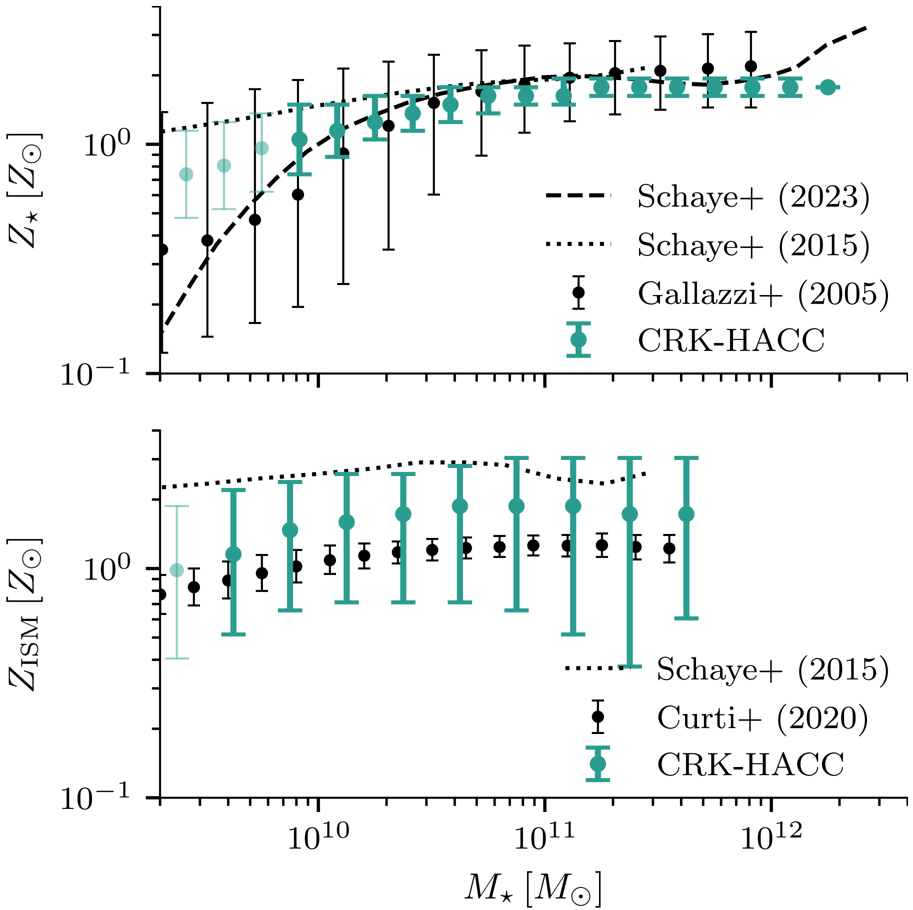


nIFTy Cluster Comparison

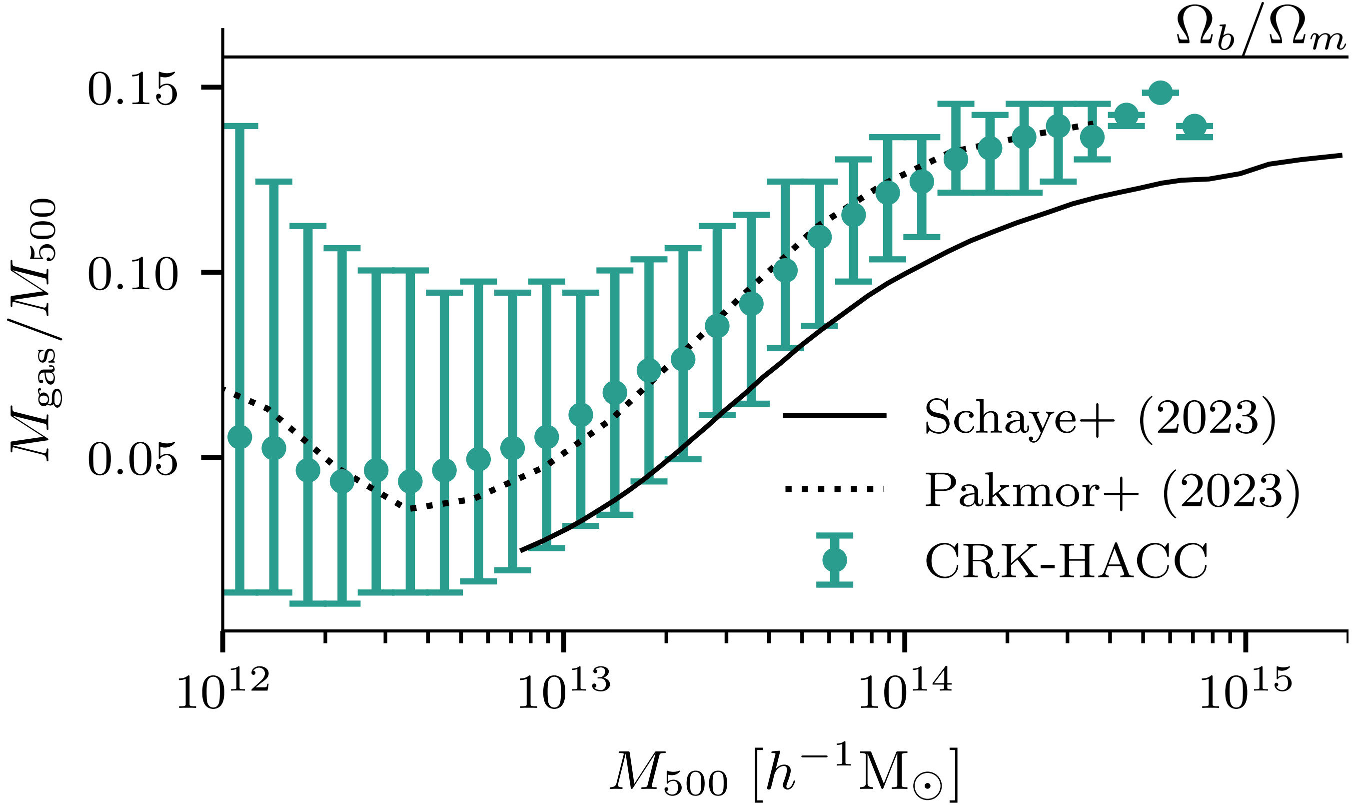


Consistent results with modern SPH, moving-mesh, and grid-based codes

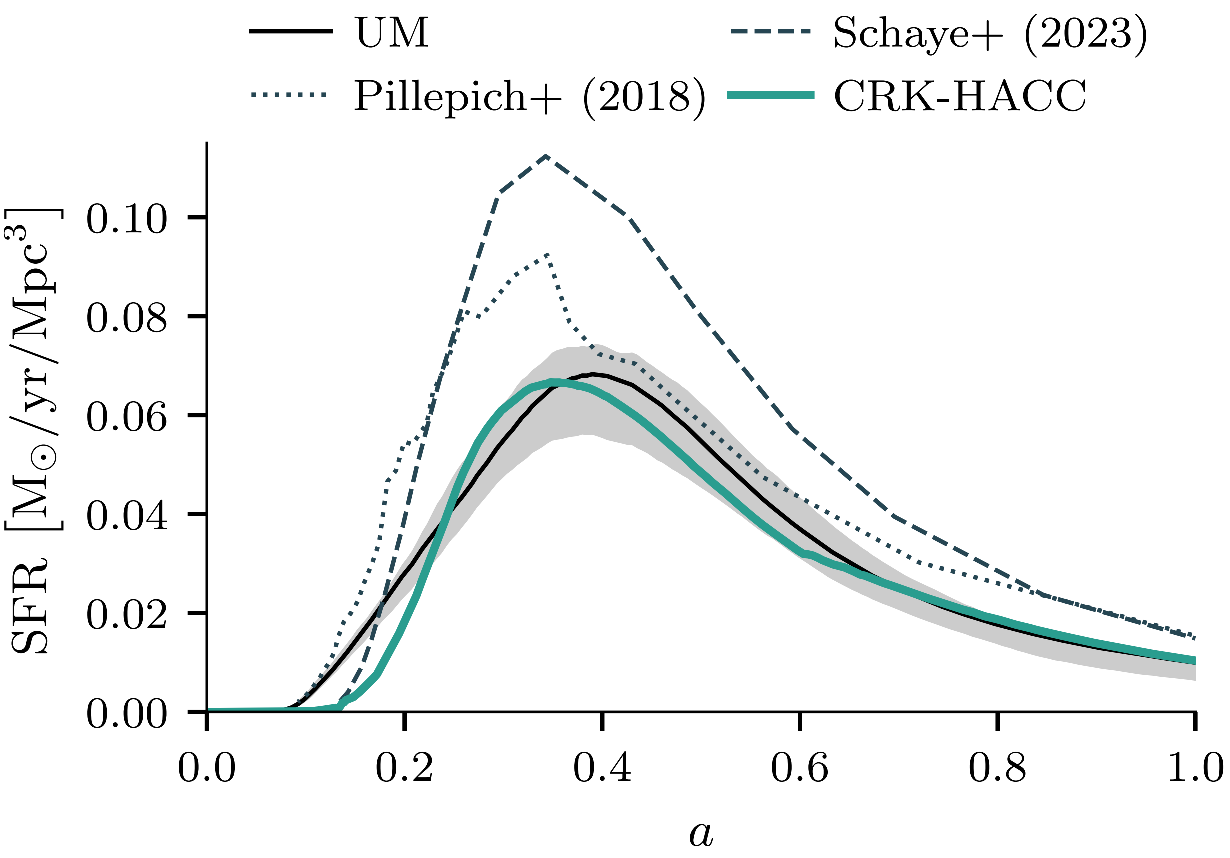
Galaxy Metallicity



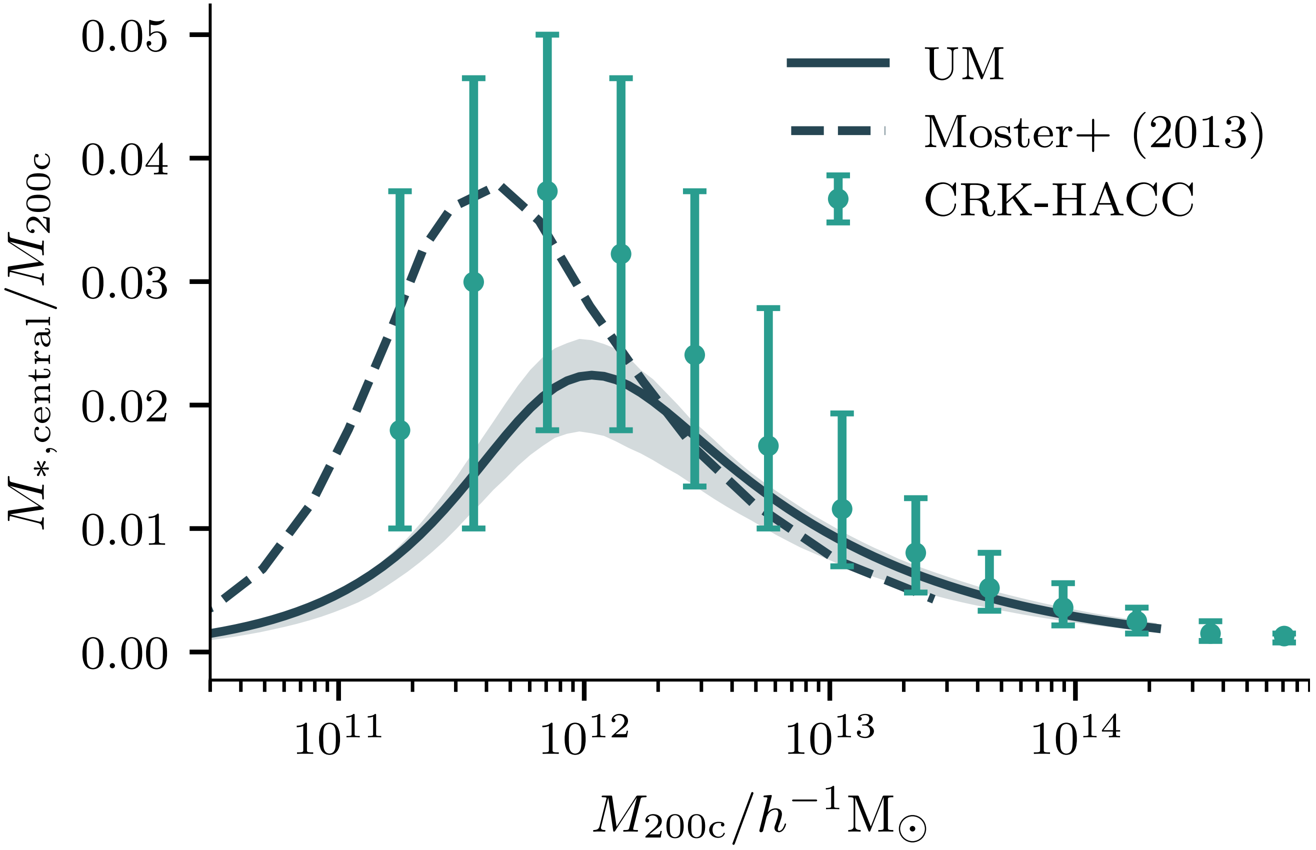
Gas Fraction



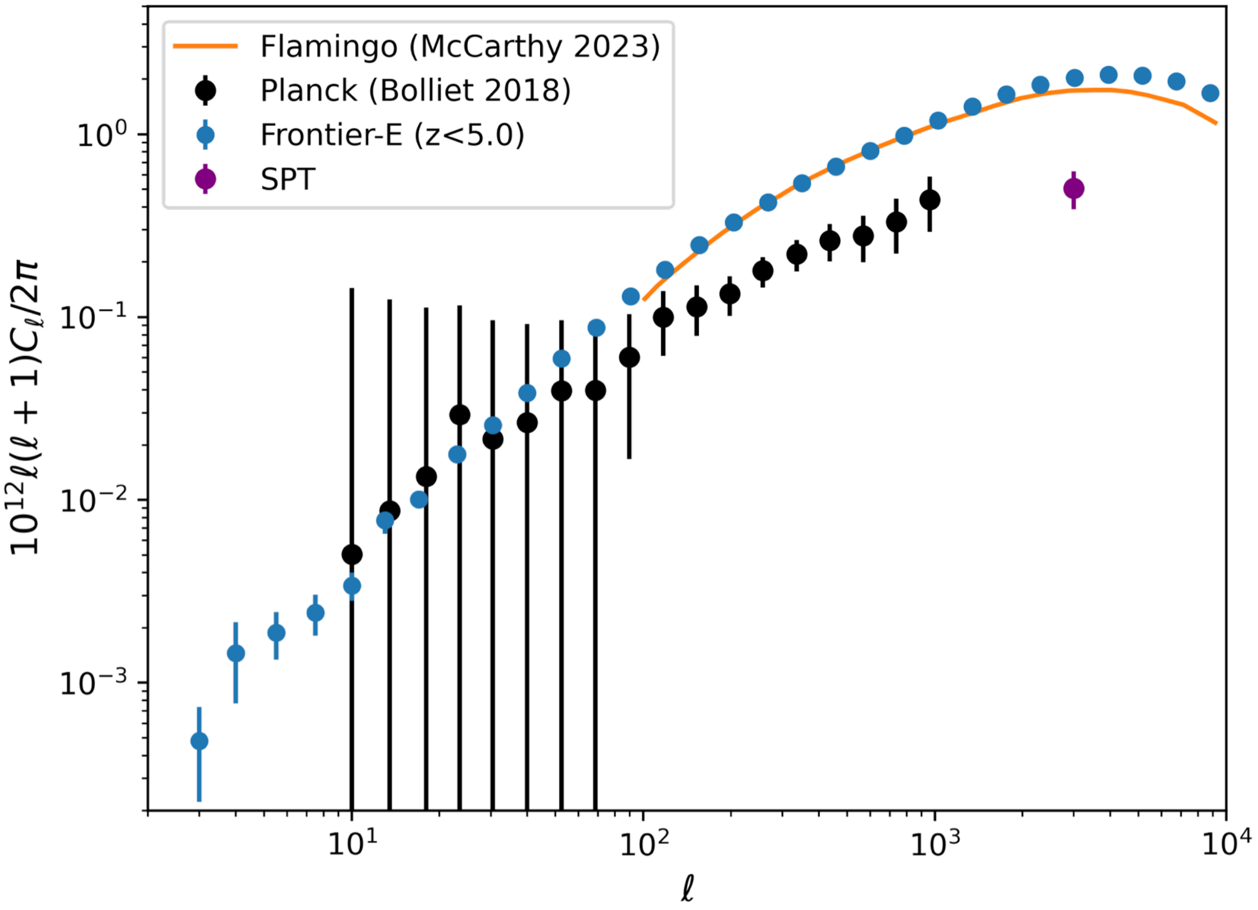
Cosmic Star Formation Rate



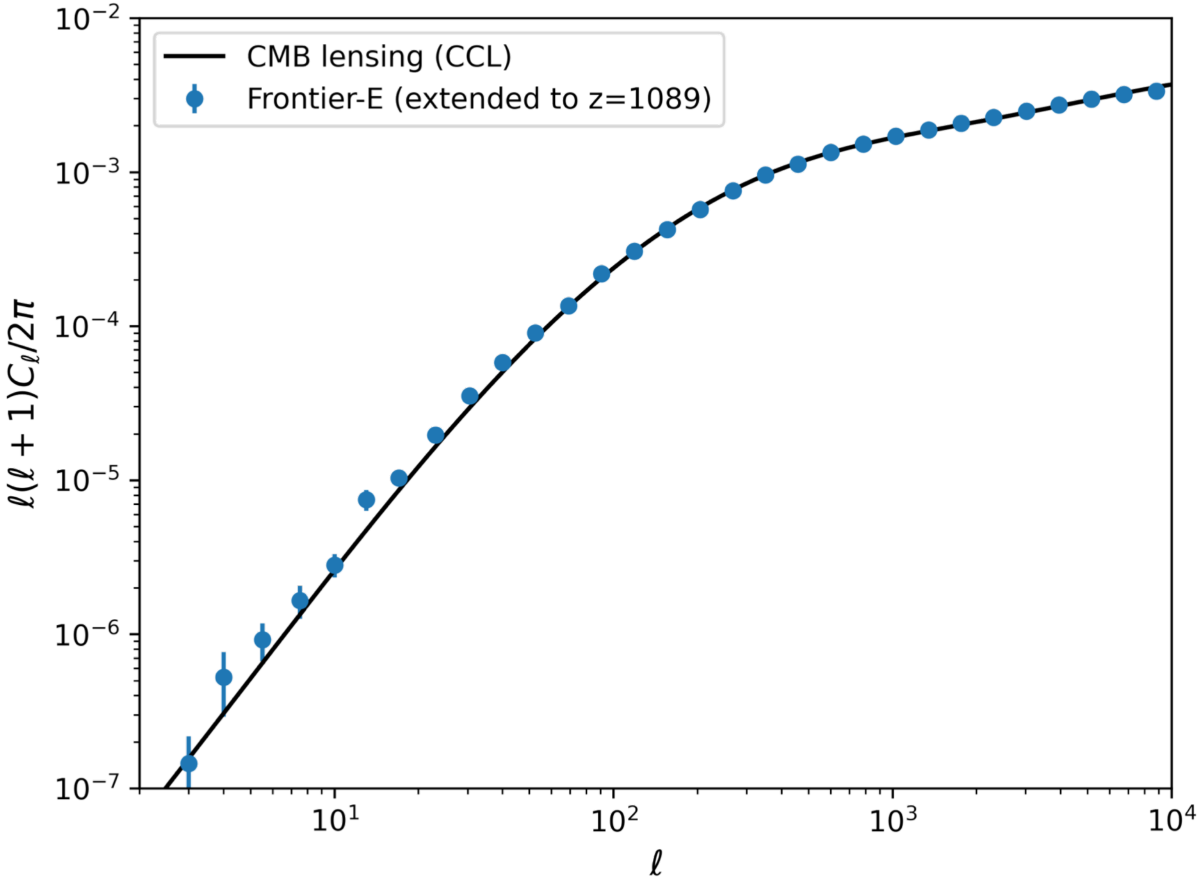
Stellar Mass - Halo Mass Relation



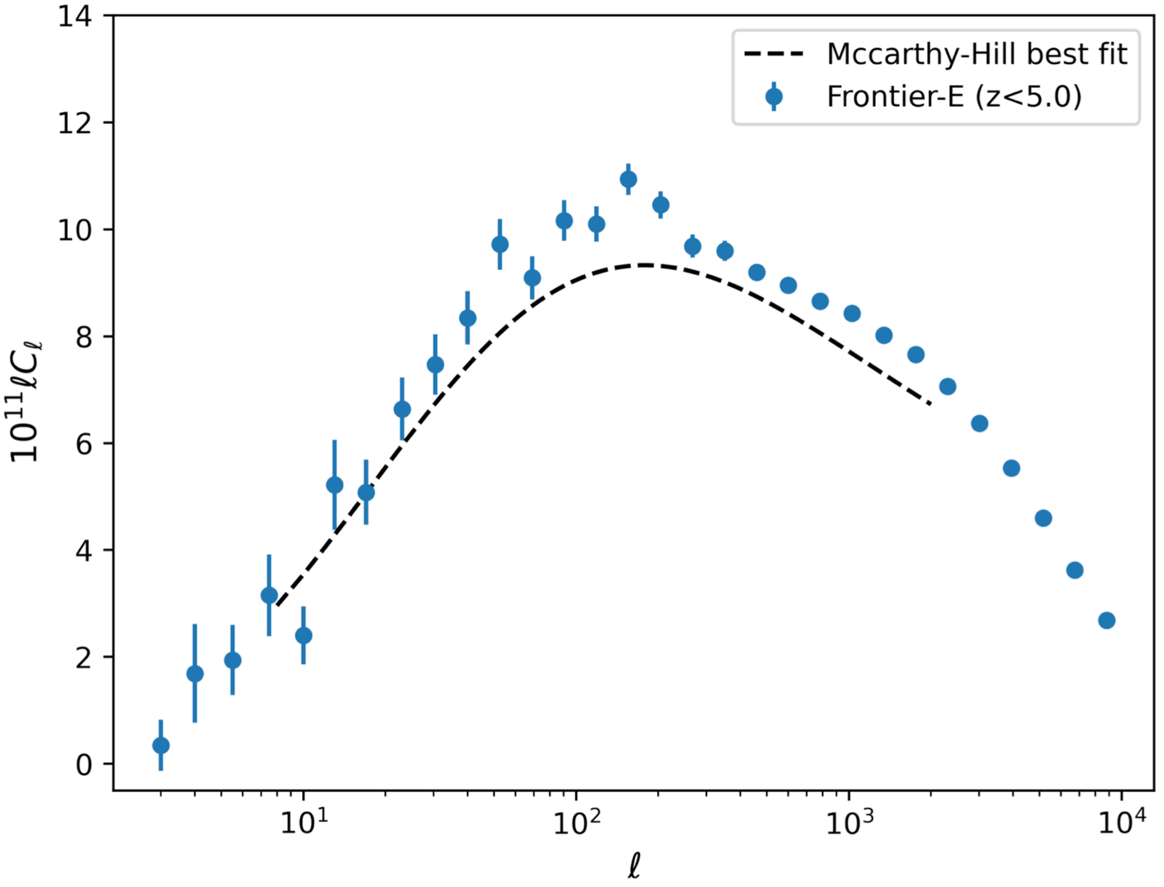
TSZ Power Spectrum



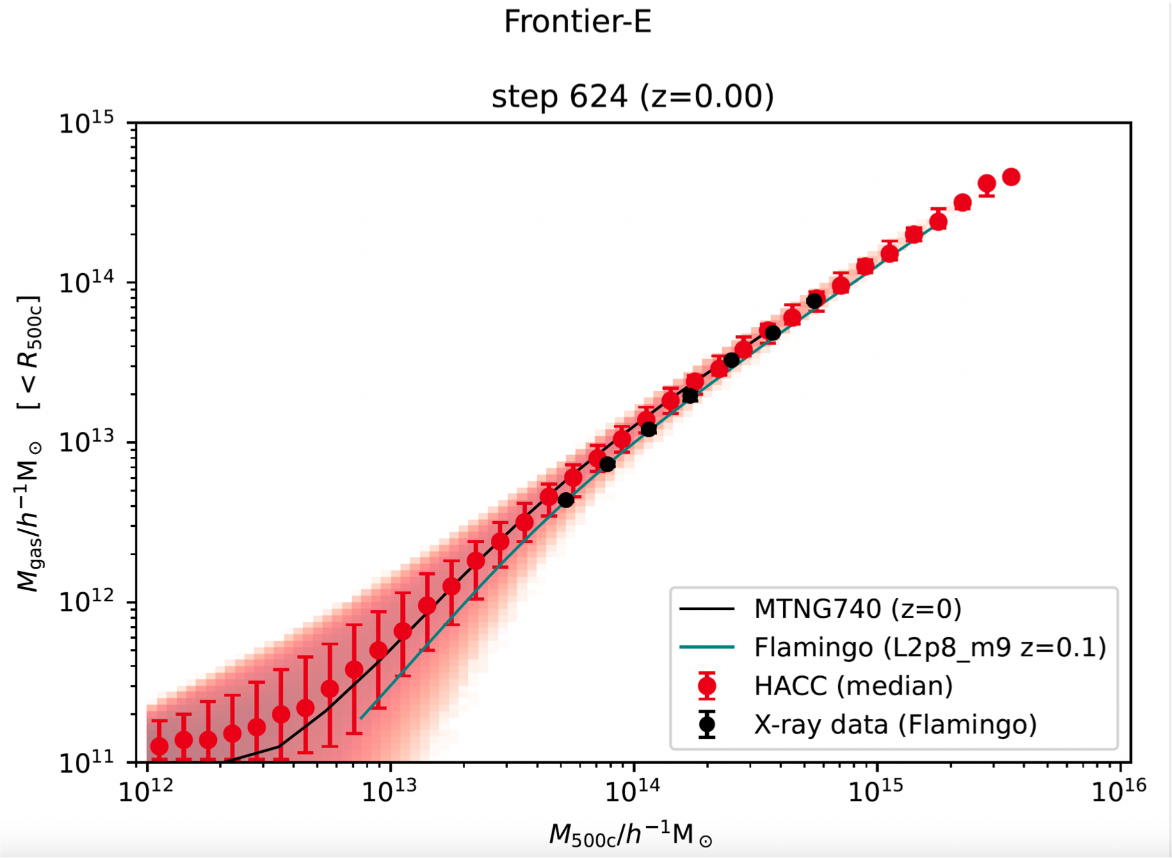
CMB Lensing


CMB lensing (CCL)
Frontier-E (extended to z=1089)
$\ell(\ell+1)C_\ell/2\pi$
ℓ
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{1}
10^{2}
10^{3}
10^{4}

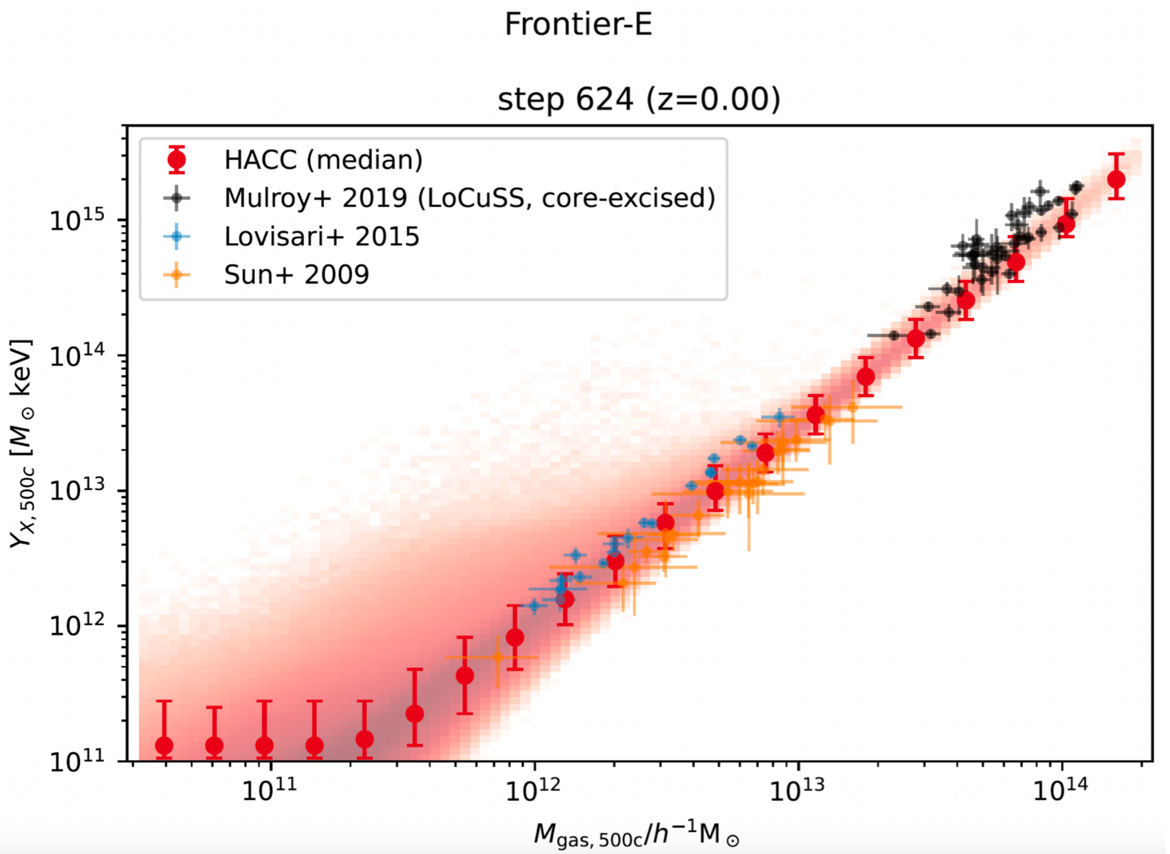
TSZ x CMB Lensing



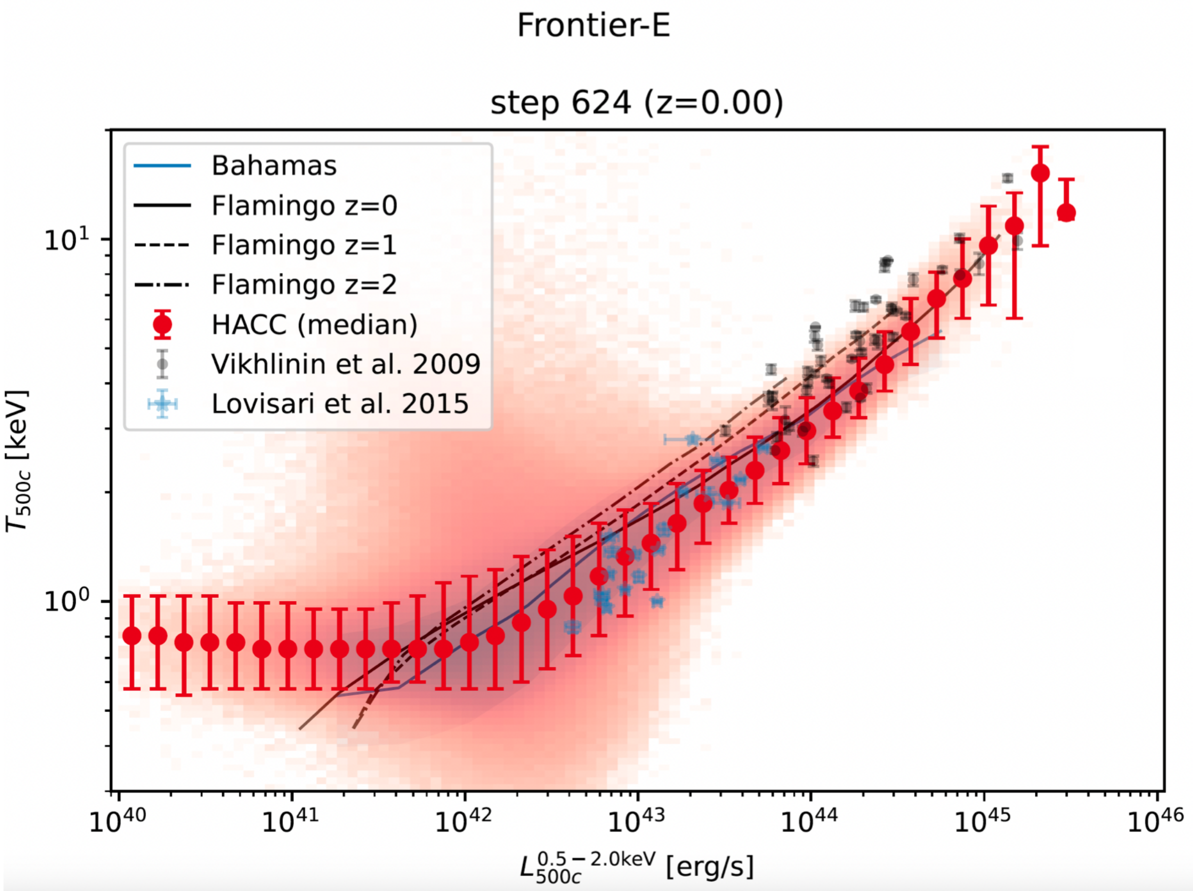
Scaling Relations — M_{gas}-M_{500}



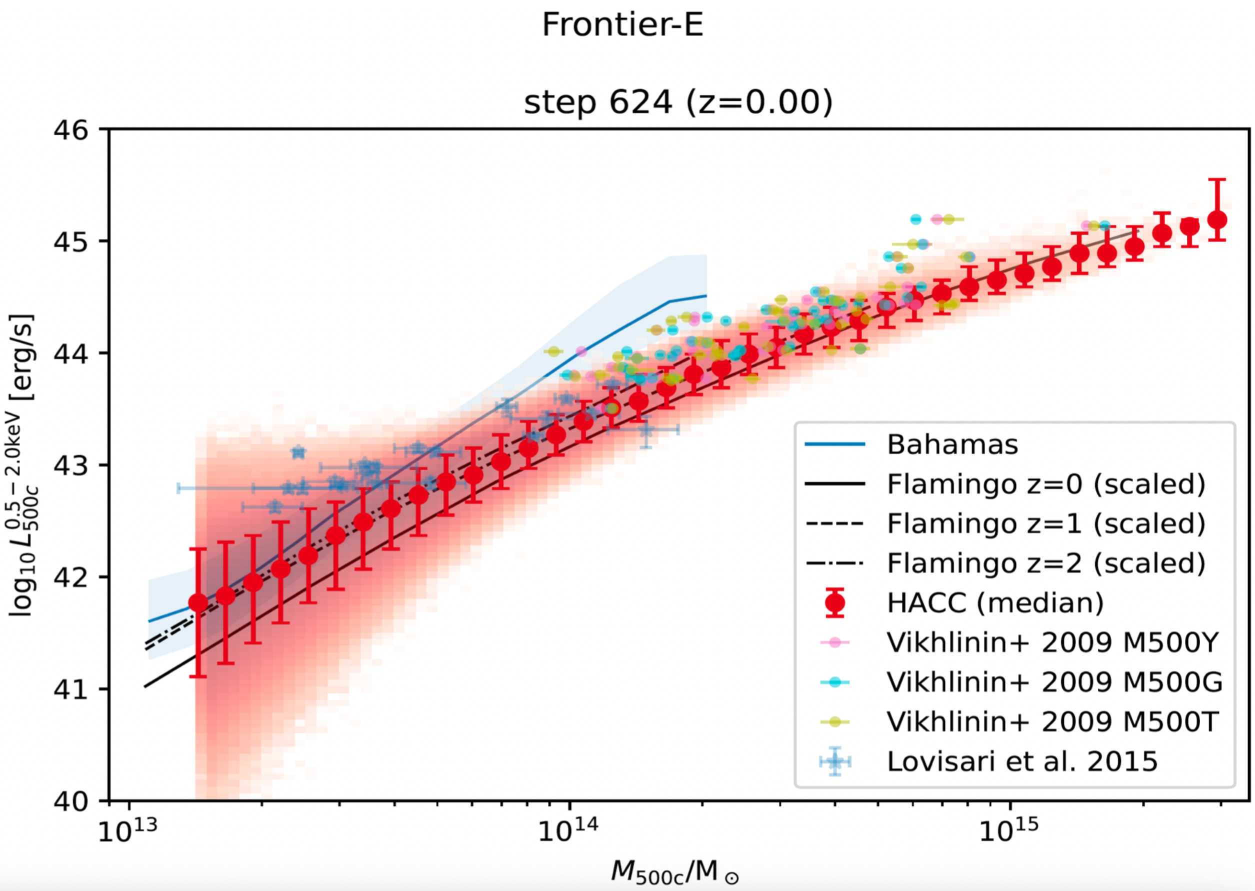
Scaling Relations — Y_x-M_{gas}



Scaling Relations — T_{500}-L_{500}

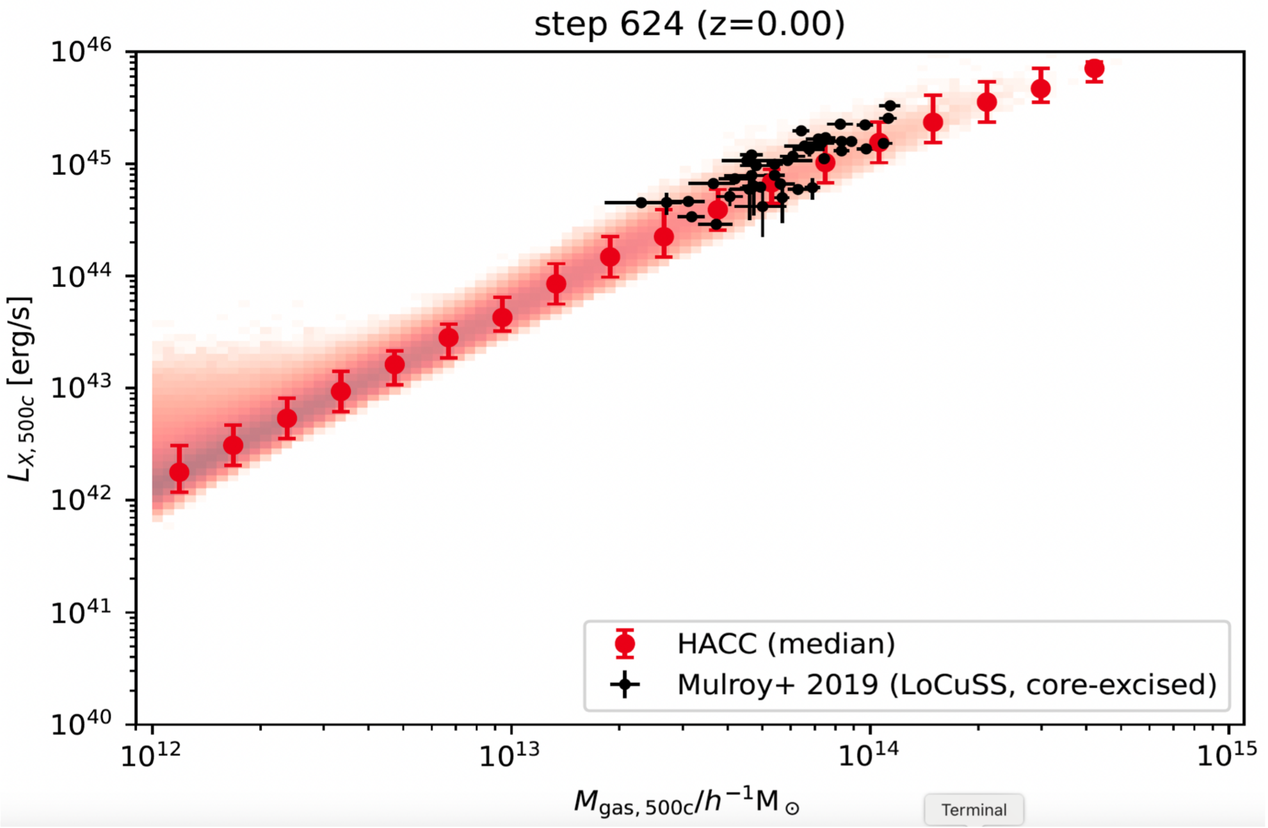


Scaling Relations — L_{500}-M_{500}

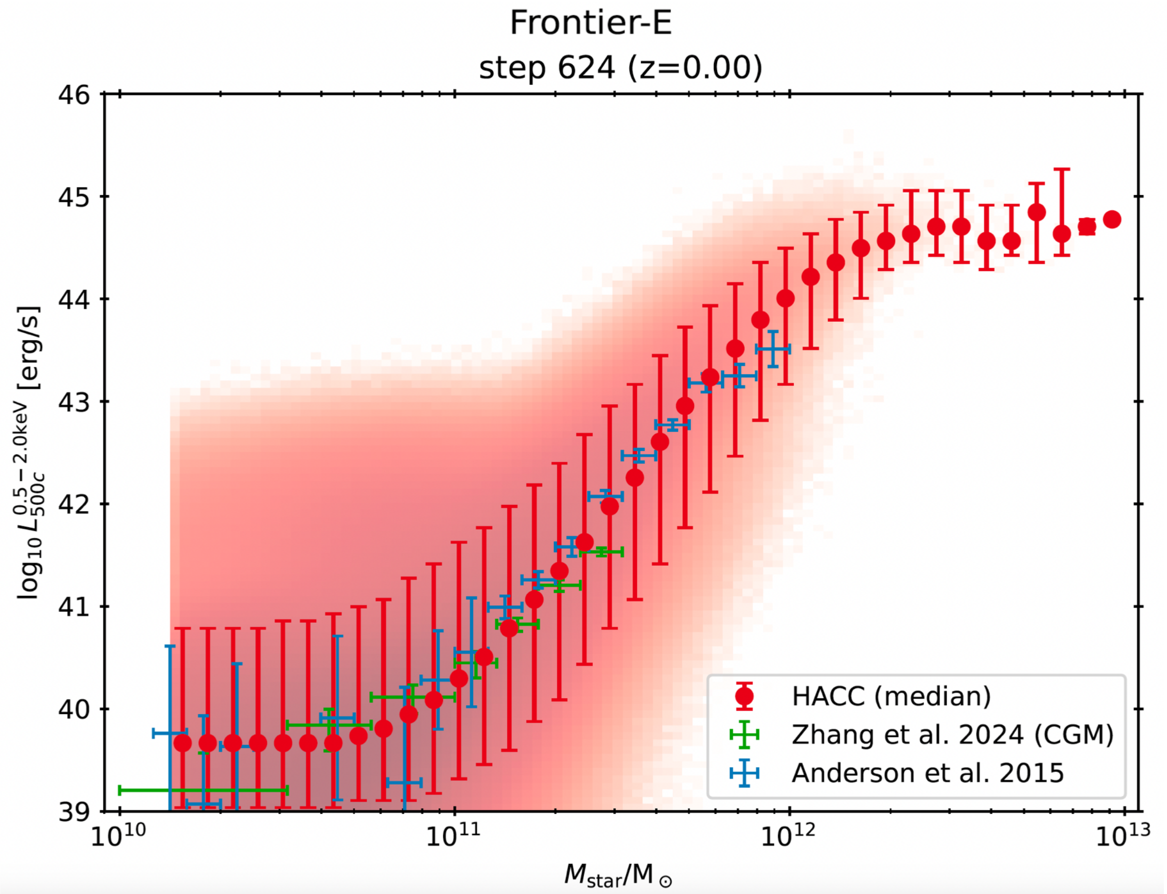


Scaling Relations — L_{500}-M_{gas}

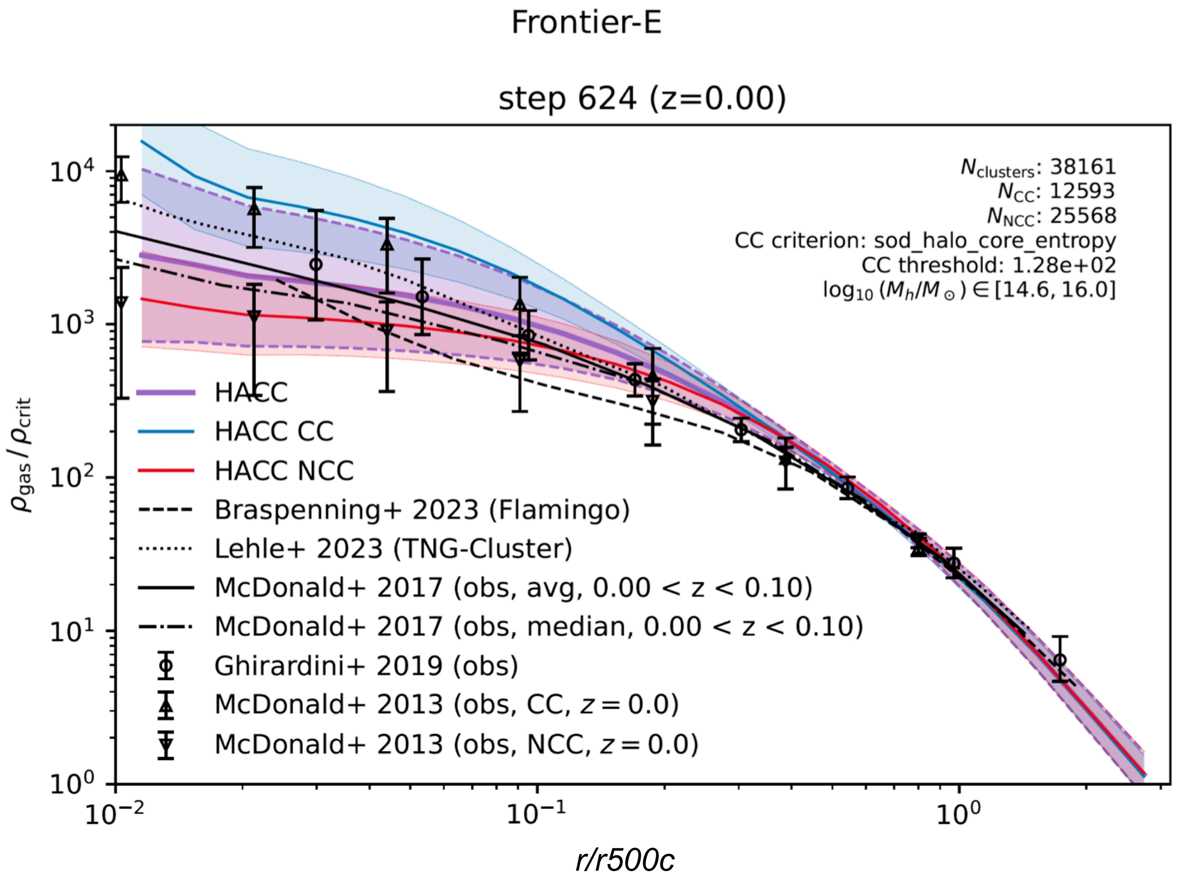
Frontier-E



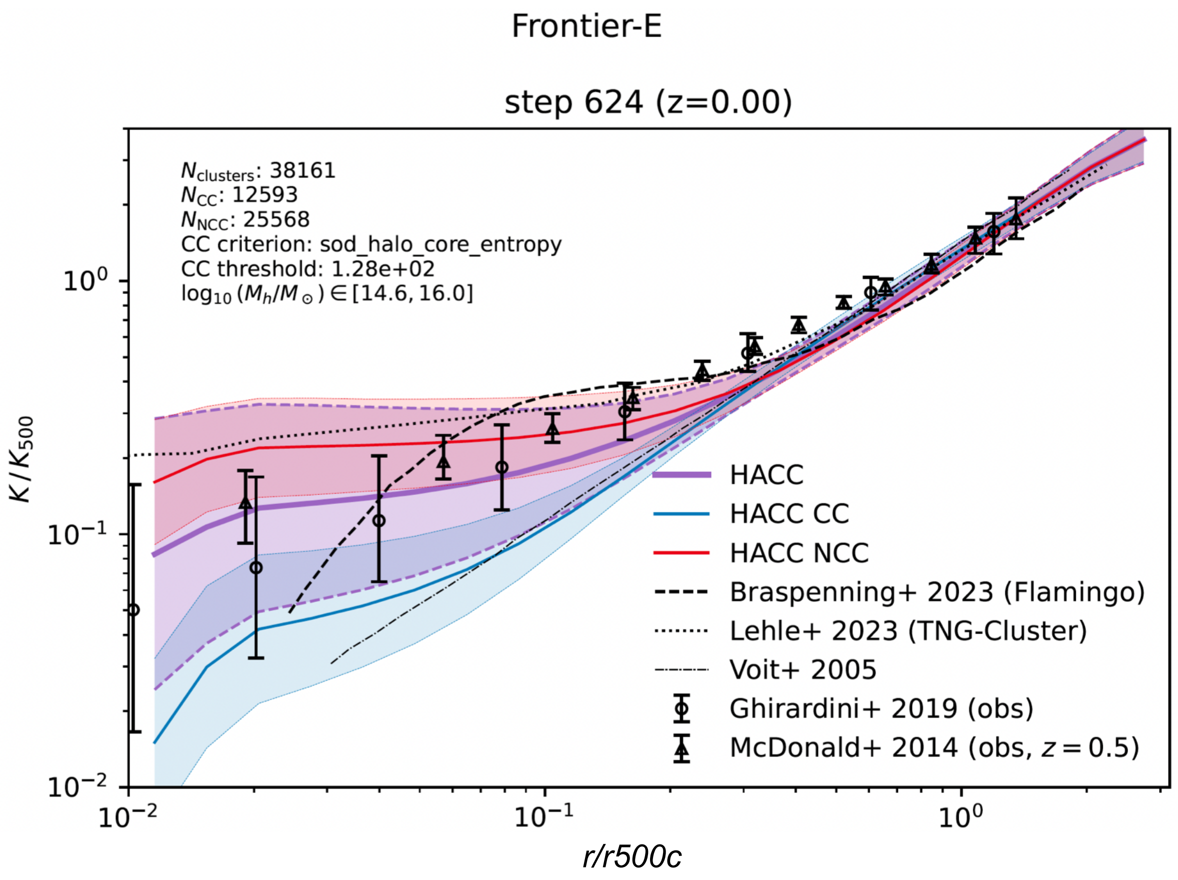
Scaling Relations — L_{500}-M_{star}



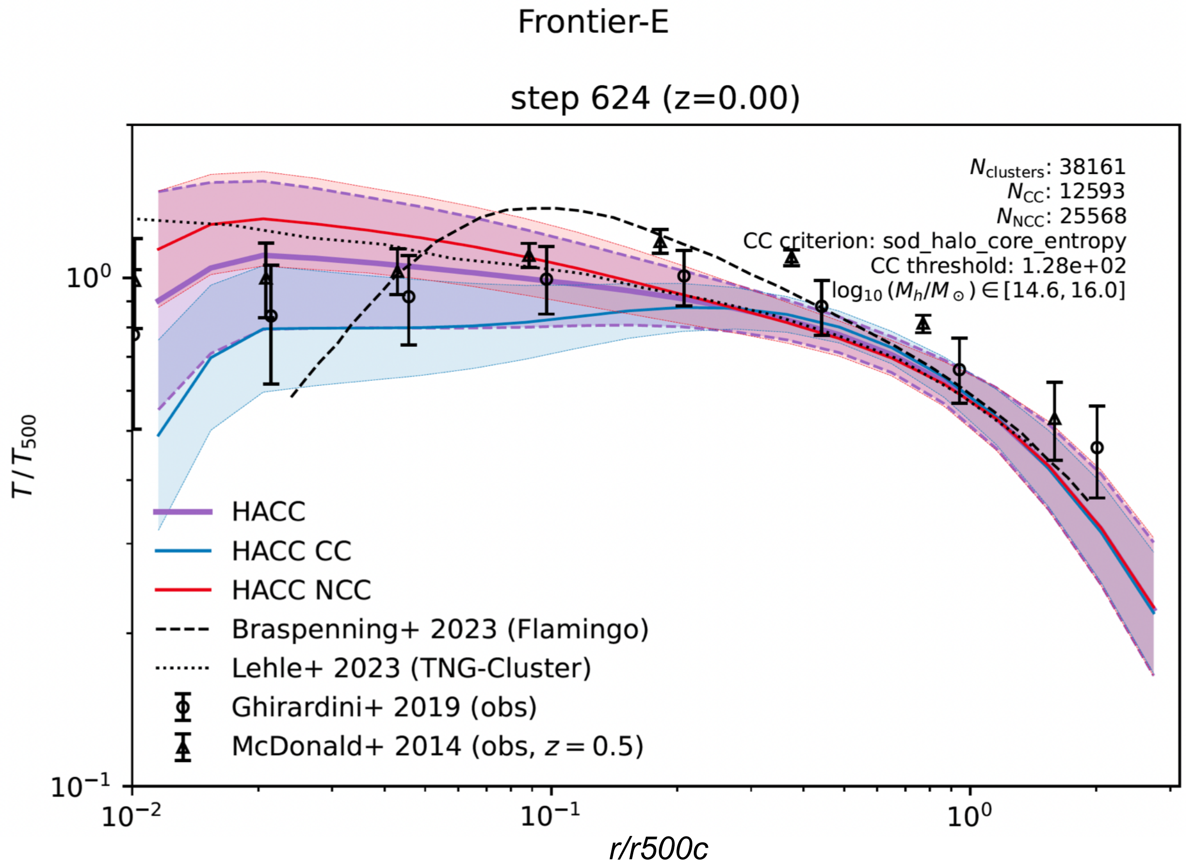
Cluster Profiles — Gas Density



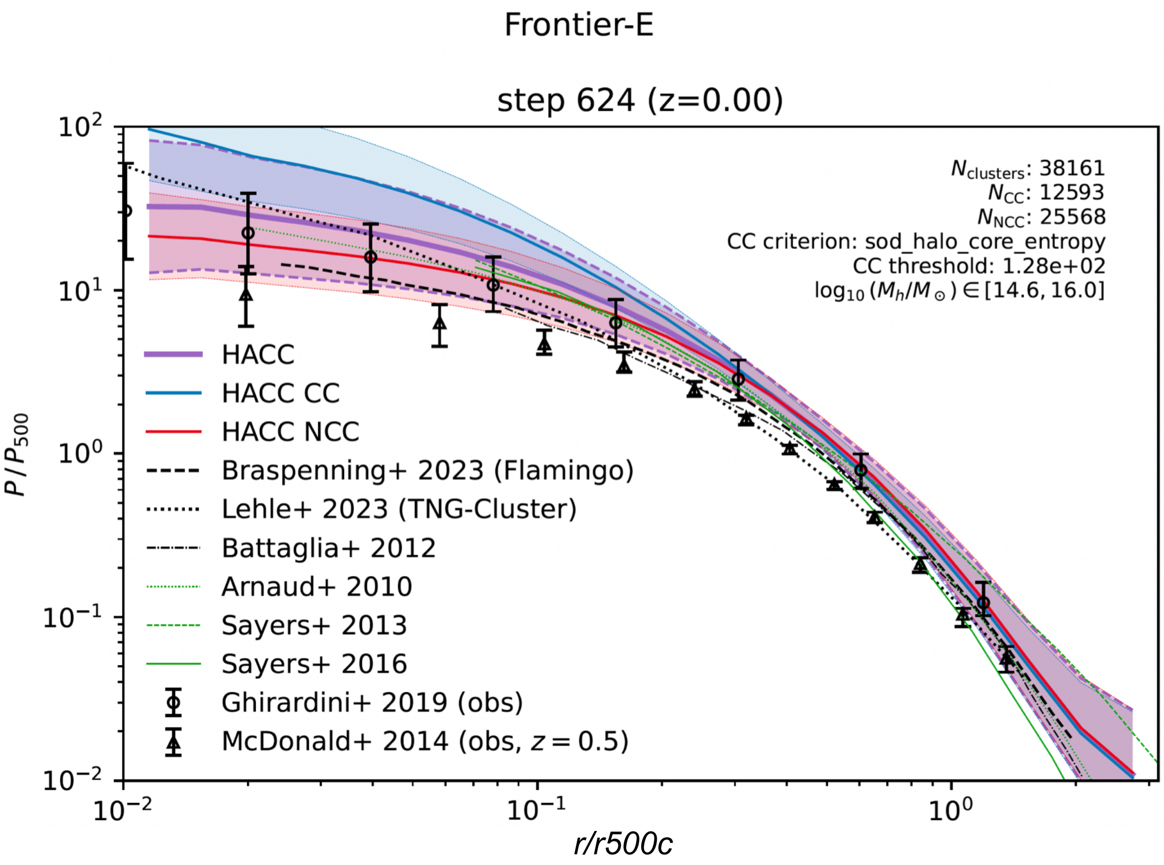
Cluster Profiles — Gas Entropy



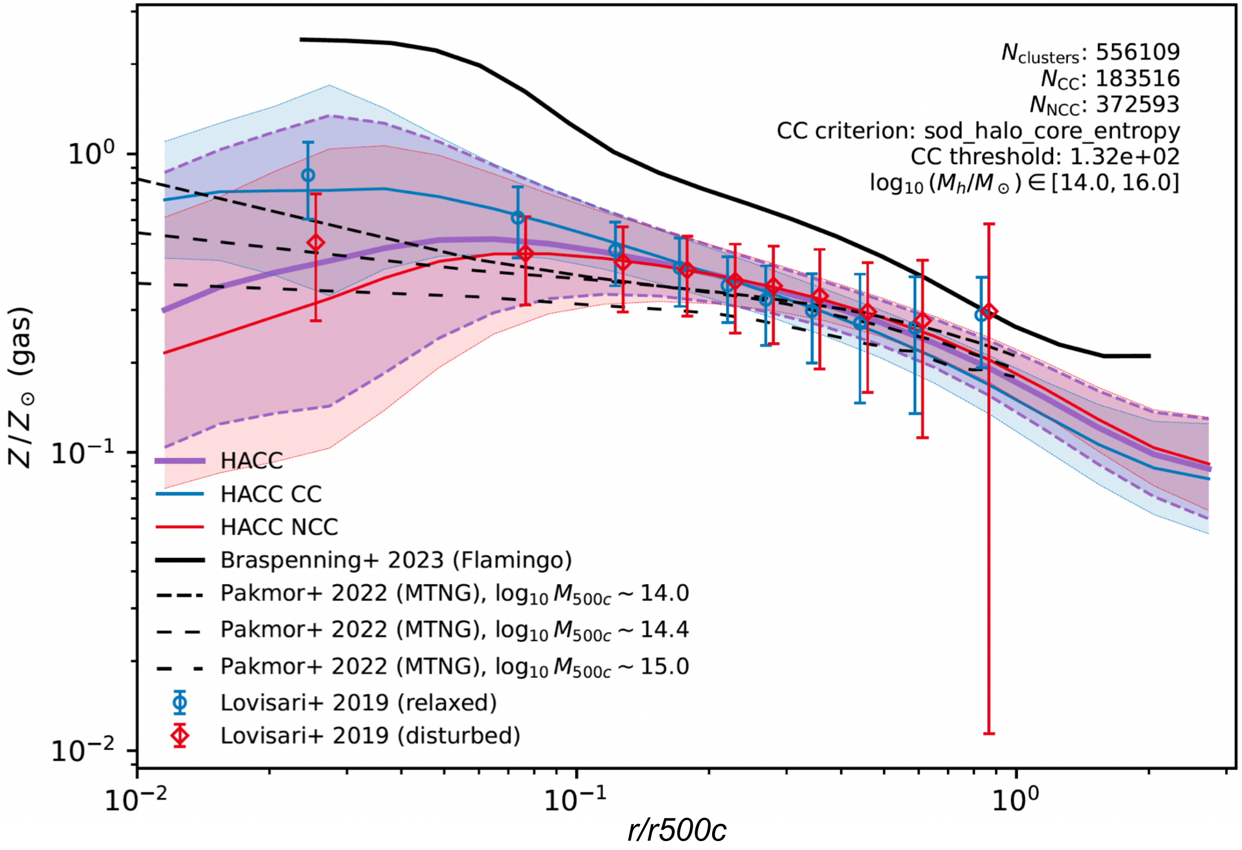
Cluster Profiles — Gas temperature



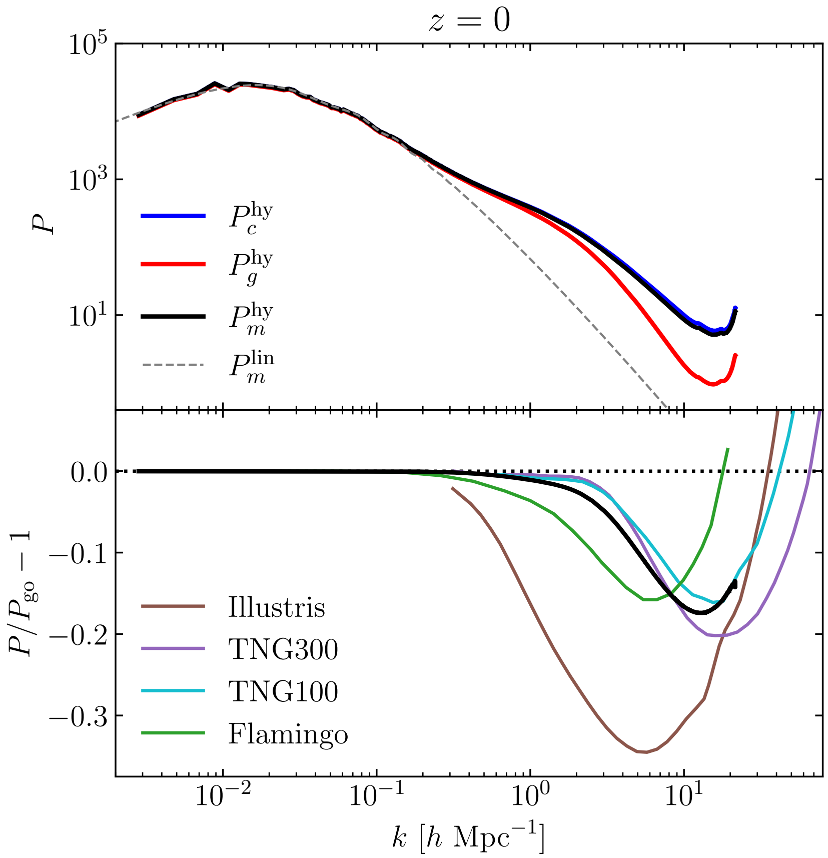
Cluster Profiles — Gas Pressure



Cluster Profiles — Metallicity

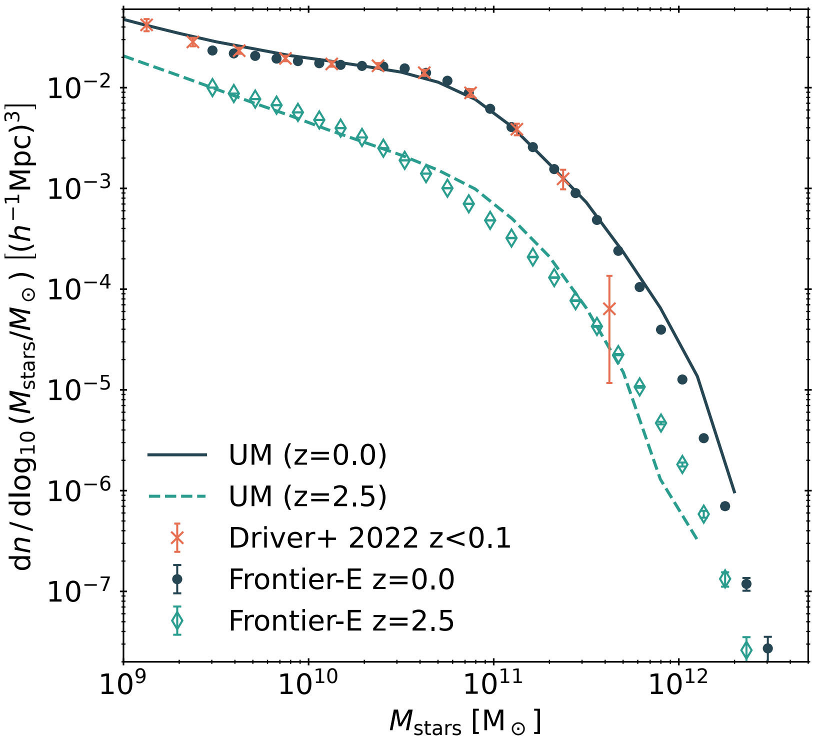


Power Spectrum



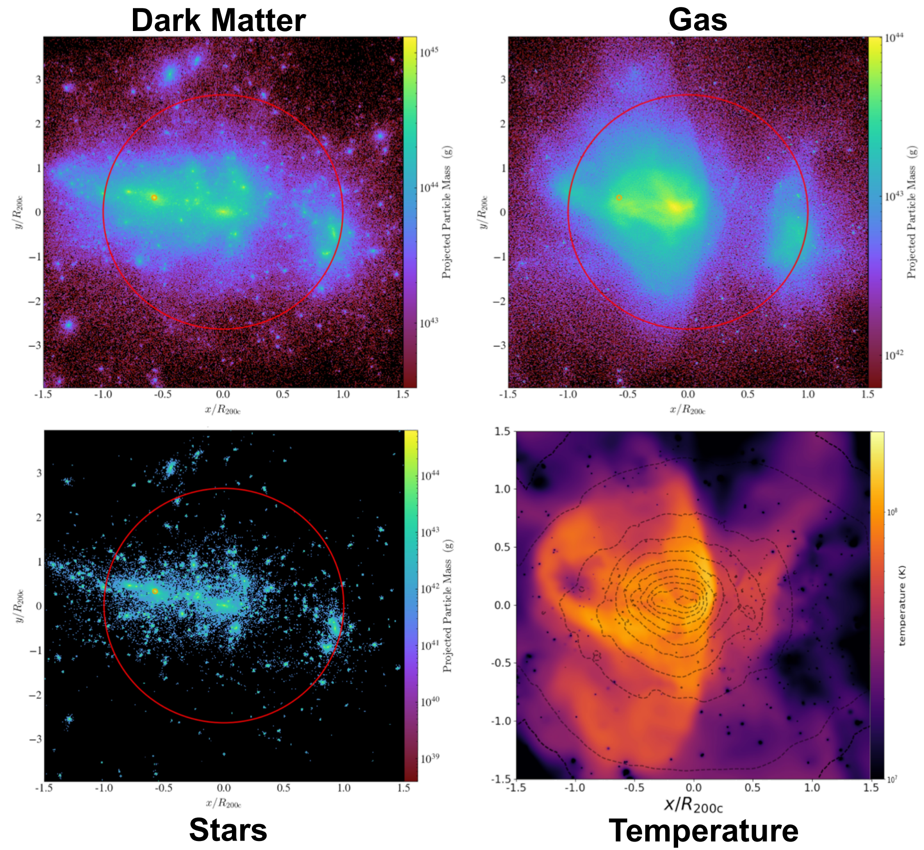
Power Spectrum Comparisons

GSMF



Calibrated Galaxy Stellar Mass Function

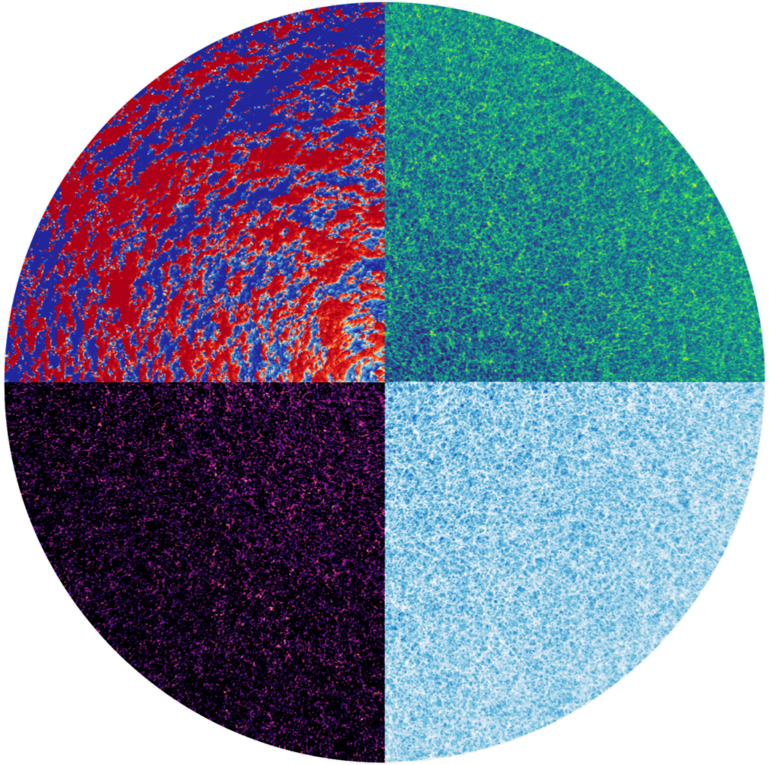
Rare Cluster Example



The large volume enables the identification and study of rare, dynamically complex systems

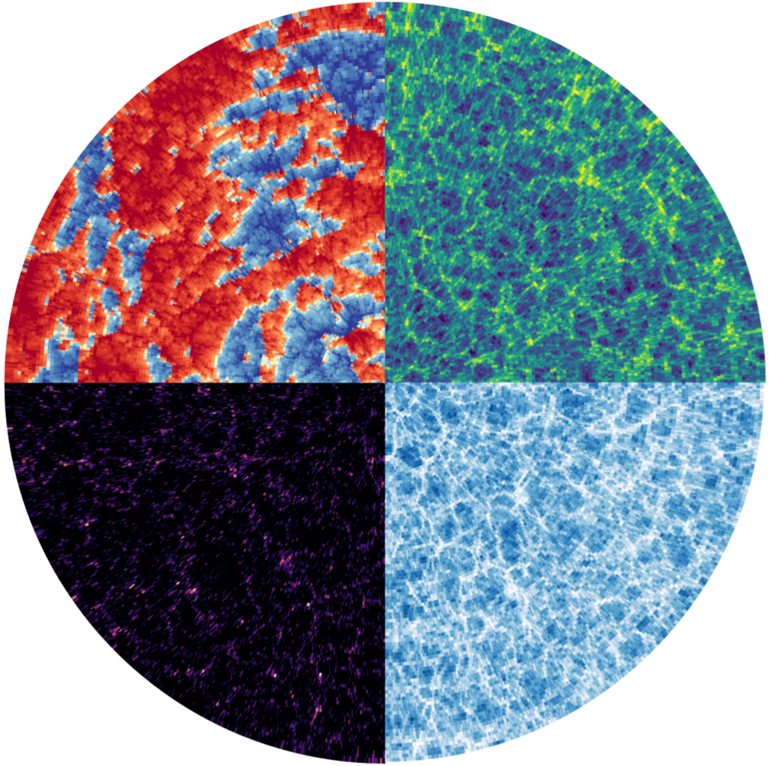
Full-sky maps

z = 1.0



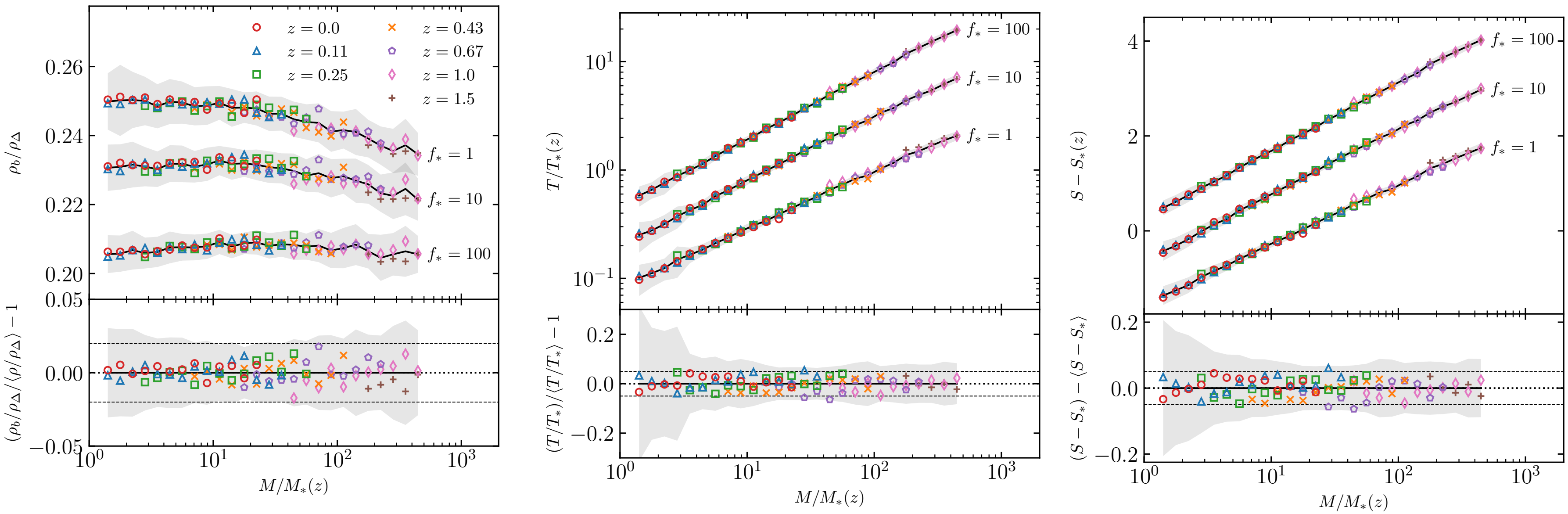
Full-sky maps showing, clockwise from top left: kSZ, tSZ, X-ray, and density.

z = 0.25

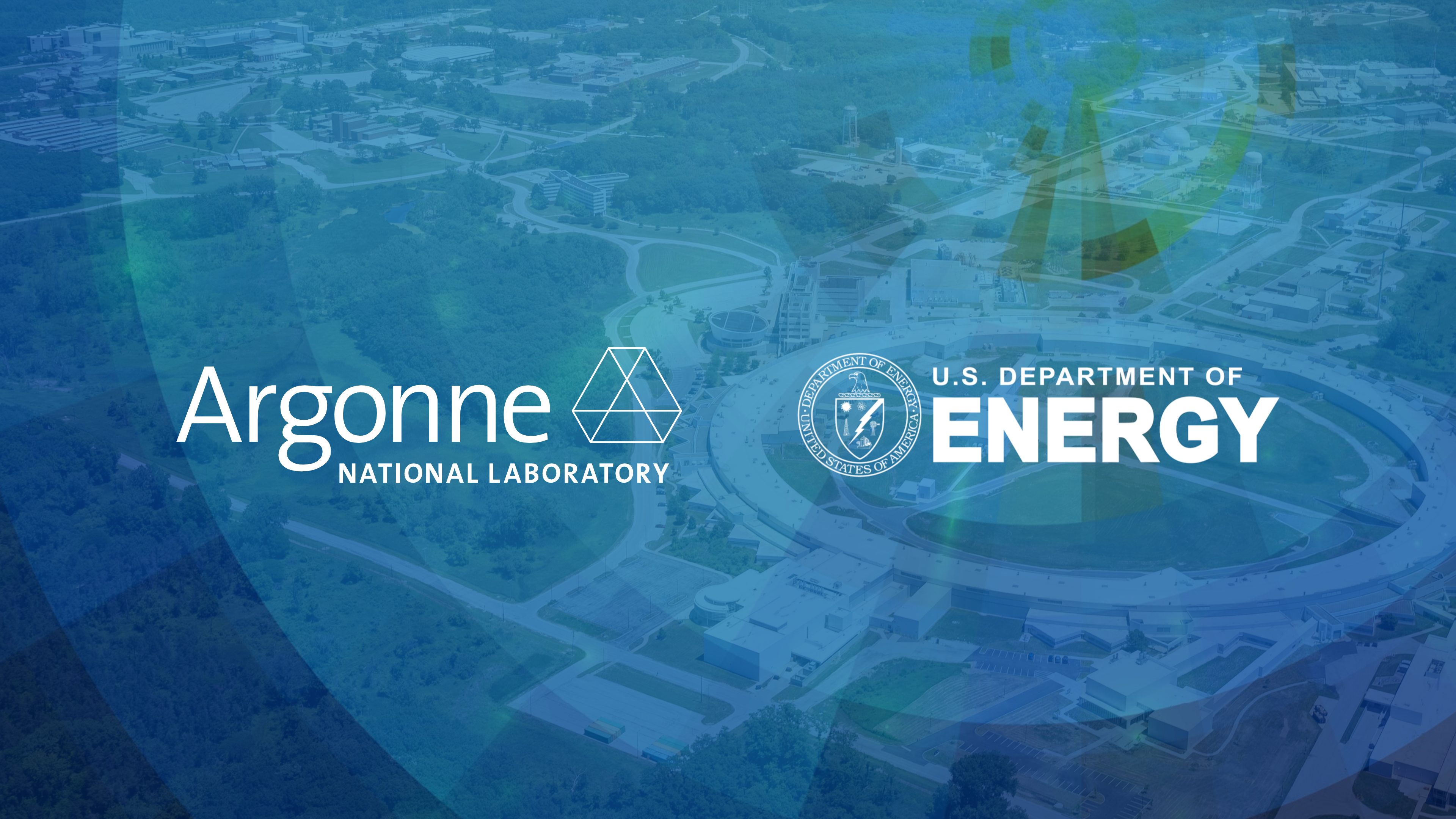


Full-sky maps showing, clockwise from top left: kSZ, tSZ, X-ray, and density.

Self-similarity with Cooling



Assuming a power-law cooling source can check self-similarity


Argonne
NATIONAL LABORATORY
U.S. DEPARTMENT OF
ENERGY